(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification⁷: C12N

(21) International Application Number: PCT/US03/17772

(22) International Filing Date: 5 June 2003 (05.06.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/386,651   6 June 2002 (06.06.2002)   US

(71) Applicant (for all designated States except US): AVALON PHARMACEUTICALS, INC. [US/US]; 20358 Seneca Meadows Parkway, Germantown, MD 20876 (US).

(72) Inventor; and
(75) Inventor/Applicant (for US only): EBNER, Reinhard [US/US]; 9906 Shelburne Terrace #316, Gaithersburg, MD 20878 (US).

(74) Agents: GRANT, Alan, J. et al.; Bain, Gilfillan, Cecchi, Stewart & Olstein, 6 Becker Farm Road, Roseland, NJ 07068 (US).

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: CANCER-LINKED GENE AS TARGET FOR CHEMOTHERAPY

(57) Abstract: Cancer-linked gene sequences, and derived amino acid sequences, are disclosed along with processes for assaying potential antitumor agents based on their modulation of the expression of these cancer-linked genes. Also disclosed are antibodies that react with the disclosed polypeptides and methods of using the antibodies to treat cancerous conditions, such as by using the antibody to target cancerous cells in vivo for purposes of delivering therapeutic agents thereto. Also described are methods of diagnosing using the gene sequences as well as the polypeptide sequences, especially for diagnosing prostate cancer.

CANCER-LINKED GENE AS TARGET FOR CHEMOTHERAPY

This application claims priority of U.S. Provisional Patent Application 60/386,651, filed 6 June 2002, the disclosure of which is hereby incorporated by reference in its entirety.

FIELD OF THE INVENTION

The present invention relates to methods of screening cancer-linked genes and expression products for involvement in the cancer initiation and facilitation process as a means of cancer diagnosis as well as the use of such genes for screening potential anti-cancer agents, including small organic compounds and other molecules, and development of therapeutic agents.

BACKGROUND OF THE INVENTION

Cancer-linked genes are valuable in that they indicate genetic differences between cancer cells and normal cells, such as where a gene is expressed in a cancer cell but not in a non-cancer cell, or where said gene is over-expressed or expressed at a higher level in a cancer as opposed to normal or non-cancer cell. In addition, the expression of such a gene in a normal cell but not in a cancer cell, especially of the same type of tissue, can indicate important functions in the cancerous process. For example, screening assays for novel drugs are based on the response of model cell based systems *in vitro* to treatment with specific compounds. Such genes are also useful in the diagnosis of cancer and the identification of a cell as cancerous.

Gene activity is readily measured by measuring the rate of production of gene products, such as RNAs and polypeptides encoded by such genes. Where genes encode cell surface proteins, appearance of, or alterations in, such proteins, as cell surface markers, are an indication of neoplastic activity.

5   Some such screens rely on specific genes, such as oncogenes (or gene mutations). In accordance with the present invention, a cancer-linked gene has been identified and its putative amino acid sequence worked out. Such gene is useful in the diagnosing of cancer, the screening of anticancer agents and the treatment of cancer using such agents, especially in that these genes

10 encode polypeptides that can act as markers, such as cell surface markers, thereby providing ready targets for anti-tumor agents such as antibodies, preferably antibodies complexed to cytotoxic agents, including apoptotic agents.

15

BRIEF SUMMARY OF THE INVENTION

In accordance with the present invention, there is provided herein a

20 cancer specific gene, linked especially to prostate   cancer, or otherwise involved in the cancer initiating and facilitating process and the derived amino acid sequence thereof, including a number of different transcripts derived from said gene.

25   In one aspect, the present invention relates to a process for identifying an agent that modulates the activity of a cancer-related gene comprising:

(a) contacting a compound with a cell containing a gene that corresponds to a polynucleotide having a sequence selected from the group consisting of SEQ ID NO: 1 - 18 and under conditions promoting the

30 expression of said gene; and

(b) detecting a difference in expression of said gene relative to when said compound is not present

2

thereby identifying an agent that modulates the activity of a cancer-related gene.

In various embodiments of such a process, the cell is a cancer cell and the difference in expression is a decrease in expression. Such polynucleotides may also include those that have sequences identical to SEQ ID NO: 1 - 18.

In another aspect, the present invention relates to a process for identifying an anti-neoplastic agent comprising contacting a cell exhibiting neoplastic activity with a compound first identified as a cancer related gene modulator using an assay process disclosed herein and detecting a decrease in said neoplastic activity after said contacting compared to when said contacting does not occur. Such neoplastic activity may include accelerated cellular replication and/or metastasis, and the decrease in neoplastic activity preferably results from the death of the cell, or senescence, terminal differentiation or growth inhibition.

The present invention also relates to a process for identifying an anti-neoplastic agent comprising administering to an animal exhibiting a cancer condition an effective amount of an agent first identified according to a process of one of one of the assays disclosed according to the invention and detecting a decrease in said cancerous condition.

The present invention further relates to a process for determining the cancerous status of a cell, comprising determining an increase in the level of expression in said cell of at least one gene that corresponds to a polynucleotide having a sequence selected from the group consisting of SEQ ID NO: 1 - 18 wherein an elevated expression relative to a known non-cancerous cell indicates a cancerous state or potentially cancerous state. Such elevated expression may be due to an increased copy number.

The present invention additionally relates to an isolated polypeptide, encoded by one of the polynucleotide transcripts disclosed herein, comprising an amino acid sequence homologous to an amino acid sequence selected from the group consisting of SEQ ID NO: 19 - 26 wherein any difference

5   between said amino acid sequence and the sequence of SEQ ID NO: 19 - 26 is due solely to conservative amino acid substitutions and wherein said isolated polypeptide comprises at least one immunogenic fragment. In a preferred embodiment, the present invention encompasses an isolated polypeptide comprising an amino acid sequence homologous to an amino

10   acid sequence selected from the group consisting of SEQ ID NO: 19 - 26.

The present invention also relates to an antibody that reacts with a polypeptide as disclosed herein, preferably a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO: 19 -

15   26. Such an antibody may be polyclonal, monoclonal, recombinant or synthetic in origin.

In one such embodiment, said antibody is associated, either covalently or non-covalently, with a cytotoxic agent, for example, an

20   apoptotic   agent.   Thus,   the   present   invention   relates   to   an immunoconjugate comprising an antibody of the invention and a cytotoxic agent.

The present invention also relates to a process for treating cancer

25   comprising contacting a cancerous cell with an agent having activity against an expression product encoded by a gene sequence selected from the group consisting of SEQ ID NO: 1 - 18. In one such embodiment, the cancerous cell is contacted *in vivo*. In another such embodiment, said agent has affinity for said expression product. In a preferred embodiment, such agent is an

30   antibody disclosed herein, such as an antibody that is specific or selective for, or otherwise reacts with, a polypeptide of the invention. In a preferred

embodiment, the expression product is a polypeptide incorporating an amino acid sequence selected from SEQ ID NO: 19 - 26.

5   The present invention further encompasses an immunogenic composition comprising a polypeptide disclosed herein, as well as compositions formed using antibodies specific for these polypeptides.

    The present invention is also directed to uses of such compositions. Such uses include a method for treating cancer in an animal afflicted 10 therewith comprising administering to said animal an amount of an immunogenic composition of one or more of the polypeptides disclosed herein where such amount is an amount sufficient to elicit the production of cytotoxic T lymphocytes specific for a polypeptide of the invention, preferably a polypeptide incorporating a sequence of SEQ ID NO: 19 - 26. In a preferred 15 embodiment, the animal to be so treated is a human patient.

DEFINITIONS

20   As used herein, the terms "portion," "segment," and "fragment," when used in relation to polypeptides, refer to a continuous sequence of residues, such as amino acid residues, which sequence forms a subset of a larger sequence. For example, if a polypeptide were subjected to treatment with any of the common endopeptidases, such as trypsin or chymotrypsin, the oligopeptides 25 resulting from such treatment would represent portions, segments or fragments of the starting polypeptide. When used in relation to a polynucleotides, such terms refer to the products produced by treatment of said polynucleotides with any of the common endonucleases.

30   As used herein, the term "isolated" means that the material is removed from its original environment (e.g., the natural environment if it is naturally occurring). It could also be produced recombinantly and subsequently purified.

For example, a naturally-occurring polynucleotide or polypeptide present in a living animal is not isolated, but the same polynucleotide or polypeptide, separated from some or all of the coexisting materials in the natural system, is isolated. Such polynucleotides, for example, those prepared recombinantly,

5   could be part of a vector and/or such polynucleotides or polypeptides could be part of a composition, and still be isolated in that such vector or composition is not part of its natural environment. In one embodiment of the present invention, such isolated, or purified, polypeptide is useful in generating antibodies for practicing the invention, or where said antibody is attached to a cytotoxic or

10   cytolytic agent, such as an apoptotic agent.

The term "percent identity" or "percent identical," when referring to a sequence, means that a sequence is compared to a claimed or described sequence after alignment of the sequence to be compared (the "Compared

15   Sequence") with the described or claimed sequence (the "Reference Sequence"). The Percent Identity is then determined according to the following formula:

Percent Identity = $100 \, [1-(C/R)]$

20

wherein C is the number of differences between the Reference Sequence and the Compared Sequence over the length of alignment between the Reference Sequence and the Compared Sequence wherein (i) each base or amino acid in the Reference Sequence that does not have a corresponding aligned base or

25   amino acid in the Compared Sequence and (ii) each gap in the Reference Sequence and (iii) each aligned base or amino acid in the Reference Sequence that is different from an aligned base or amino acid in the Compared Sequence, constitutes a difference; and R is the number of bases or amino acids in the Reference Sequence over the length of the alignment with the Compared

30   Sequence with any gap created in the Reference Sequence also being counted as a base or amino acid.

If an alignment exists between the Compared Sequence and the Reference Sequence for which the percent identity as calculated above is about equal to or greater than a specified minimum Percent Identity then the Compared Sequence has the specified minimum percent identity to the

5   Reference Sequence even though alignments may exist in which the hereinabove calculated Percent Identity is less than the specified Percent Identity.

As known in the art "similarity" between two polypeptides is determined

10 by comparing the amino acid sequence and its conserved amino acid substitutes of one polypeptide to the sequence of a second polypeptide.

In accordance with the present invention, the term "DNA segment" or "DNA sequence" refers to a DNA polymer, in the form of a separate fragment

15 or as a component of a larger DNA construct, which has been derived from DNA isolated at least once in substantially pure form, i.e., free of contaminating endogenous materials and in a quantity or concentration enabling identification, manipulation, and recovery of the segment and its component nucleotide sequences by standard biochemical methods, for

20 example, using a cloning vector. Such segments are provided in the form of an open reading frame uninterrupted by internal nontranslated sequences, or introns, which are typically present in eukaryotic genes. Sequences of non-translated DNA may be present downstream from the open reading frame, where the same do not interfere with manipulation or expression of the coding

25 regions.

The term "coding region" refers to that portion of a gene which either naturally or normally codes for the expression product of that gene in its natural genomic environment, i.e., the region coding *in vivo* for the native

30 expression product of the gene. The coding region can be from a normal, mutated or altered gene, or can even be from a DNA sequence, or gene, wholly synthesized in the laboratory using methods well known to those of skill in the art of DNA synthesis.

In accordance with the present invention, the term "nucleotide sequence" refers to a heteropolymer of deoxyribonucleotides. Generally, DNA segments encoding the proteins provided by this invention are assembled

5 from cDNA fragments and short oligonucleotide linkers, or from a series of oligonucleotides, to provide a synthetic gene which is capable of being expressed in a recombinant transcriptional unit comprising regulatory elements derived from a microbial, eukaryotic or viral operon.

10 The term "expression product" means that polypeptide or protein that is the natural translation product of the gene and any nucleic acid sequence coding equivalents resulting from genetic code degeneracy and thus coding for the same amino acid(s).

15 The term "active fragment," when referring to a coding sequence, means a portion comprising less than the complete coding region whose expression product retains essentially the same biological function or activity as the expression product of the complete coding region.

20 The term "primer" means a short nucleic acid sequence that is paired with one strand of DNA and provides a free 3'-OH end at which a DNA polymerase starts synthesis of a deoxyribonucleotide chain.

The term "promoter" means a region of DNA involved in binding of RNA

25 polymerase to initiate transcription. The term "enhancer" refers to a region of DNA that, when present and active, has the effect of increasing expression of a different DNA sequence that is being expressed, thereby increasing the amount of expression product formed from said different DNA sequence.

30 The term "open reading frame (ORF)" means a series of triplets coding for amino acids without any termination codons and is a sequence (potentially) translatable into protein.

As used herein, reference to a "DNA sequence" includes both single stranded and double stranded DNA. Thus, the specific sequence, unless the context indicates otherwise, refers to the single strand DNA of such

5   sequence, the duplex of such sequence with its complement (double stranded DNA) and the complement of such sequence.

As used herein, "corresponding genes" refers to genes that encode an RNA that is at least 90% identical, preferably at least 95% identical, most

10   preferably at least 98% identical, and especially identical, to an RNA encoded by one of the nucleotide sequences disclosed herein (i.e., SEQ ID NO: 1 - 18). Such genes will also encode the same polypeptide sequence as any of the sequences disclosed herein, preferably SEQ ID NO: 1 - 18, but may include differences in such amino acid sequences where such differences are

15   limited to conservative amino acid substitutions, such as where the same overall three dimensional structure, and thus the same antigenic character, is maintained. Thus, amino acid sequences may be within the scope of the present invention where they react with the same antibodies that react with polypeptides comprising the sequences of SEQ ID NO: 19 - 26. A

20   "corresponding gene" includes splice variants thereof.

The genes identified by the present disclosure are considered "cancer-related" genes, as this term is used herein, and include genes expressed at higher levels (due, for example, to elevated rates of expression, elevated

25   extent of expression or increased copy number) in cancer cells relative to expression of these genes in normal (i.e., non-cancerous) cells where said cancerous state or status of test cells or tissues has been determined by methods known in the art, such as by reverse transcriptase polymerase chain reaction (RT-PCR) as described in the Examples herein. In specific

30   embodiments, this relates to the genes whose sequences correspond to the sequences of SEQ ID NO: 1 - 18.

As used herein, the term "conservative amino acid substitutions" are defined herein as exchanges within one of the following five groups:

    I. Small aliphatic, nonpolar or slightly polar residues:

        Ala, Ser, Thr, Pro, Gly;

5   II. Polar, negatively charged residues and their amides:

        Asp, Asn, Glu, Gln;

    III. Polar, positively charged residues:

        His, Arg, Lys;

    IV. Large, aliphatic, nonpolar residues:

10   Met Leu, Ile, Val, Cys

    V. Large, aromatic residues:

        Phe, Tyr, Trp

15

DETAILED SUMMARY OF THE INVENTION

The present invention relates to processes for utilizing a nucleotide
20 sequence for a cancer-linked gene, polypeptides encoded by such sequences and antibodies reactive with such polypeptides in methods of treating and diagnosing cancer, preferably prostate cancer, and in carrying out screening assays for agents effective in reducing the activity of cancer-linked genes and thereby treating a cancerous condition.

25

The proteins useful in the invention are anion transport proteins and are specific for prostate tissue as well as being specific for prostate cancer. Thus, the presence of such antigens is diagnostic for prostate cancer.

30 The polypeptides disclosed herein incorporate various polynucleotide transcripts (SEQ ID NO: 1 - 18) and the derived amino acid sequence (SEQ ID NO: 19 - 26) from said transcripts are available as targets for

chemotherapeutic agents, especially anti-cancer agents, including antibodies specific for said polypeptides.

The cancer-related polynucleotide sequences disclosed herein
5   correspond to gene sequences whose expression is indicative of the cancerous status of a given cell. Such sequences are substantially identical to SEQ ID NO: 1 - 18, which represent different transcripts identified from the GenBank EST database and which exhibit cancer-specific expression. The polynucleotides of the invention are those that correspond to a sequence of
10 SEQ ID NO: 1 - 18. Such sequences have been searched within the GenBank database, especially the EST database, with the following results:

Type:	Cell-surface tumor antigen therapeutic antibody target
Tissue:	prostate
AffyFragment-ID(s):	104504
Accession(s):	AF071202
Unigene cluster-ID(s):	Hs.139336 ATP-binding cassette, sub-family C (CFTR/MRP), member 4
Chromosomal location:	13

The nucleotides and polypeptides, as gene products, used in the
30 processes of the present invention may comprise a recombinant polynucleotide or polypeptide, a natural polynucleotide or polypeptide, or a synthetic polynucleotide or polypeptide, or a recombinant polynucleotide or polypeptide.

Fragments of such polynucleotides and polypeptides as are disclosed
35 herein may also be useful in practicing the processes of the present invention. For example, a fragment, derivative or analog of the polypeptide (SEQ ID NO: 19 - 26) may be (i) one in which one or more of the amino acid residues are

substituted with a conserved or non-conserved amino acid residue (preferably a conserved amino acid residue) and such substituted amino acid residue may or may not be one encoded by the genetic code, or (ii) one in which one or more of the amino acid residues includes a substituent group, or (iii) one in which the

5   mature polypeptide is fused with another compound, such as a compound to increase the half-life of the polypeptide (for example, polyethylene glycol), or (iv) one in which the additional amino acids are fused to the mature polypeptide, such as a leader or secretory sequence or a sequence which is employed for purification of the mature polypeptide (such as a histidine hexapeptide) or a

10 proprotein sequence. Such fragments, derivatives and analogs are deemed to be within the scope of those skilled in the art from the teachings herein.

In another aspect, the present invention relates to an isolated polypeptide, including a purified polypeptide, comprising an amino acid

15 sequence at least 90% identical to the amino acid sequence of SEQ ID NO: 19 – 26. In preferred embodiments, said isolated polypeptide comprises an amino acid sequence having sequence identity of at least 95%, preferably at least about 98%, and especially is identical to, the sequence of SEQ ID NO: 19 - 26. The present invention also includes isolated active fragments of such

20 polypeptides where said fragments retain the biological activity of the polypeptide or where such active fragments are useful as specific targets for cancer treatment, prevention or diagnosis. Thus, the present invention relates to any polypeptides, or fragments thereof, with sufficient sequence homology to the sequences disclosed herein as to be useful in the production of antibodies

25 that react with (i.e., are selective or specific for) the polypeptides of SEQ ID NO: 19 - 26 so as to be useful in targeting cells that exhibit such polypeptides, or fragments, on their surfaces, thereby providing targets for such antibodies and therapeutic agents associated with such antibodies.

30 The polynucleotides and polypeptides useful in practicing the processes of the present invention may likewise be obtained in an isolated or purified form. In addition, the polypeptide disclosed herein as being useful in practicing the

processes of the invention are believed to be surface proteins present on cells, such as cancerous cells. Precisely how such cancer-linked proteins are used in the processes of the invention may thus differ depending on the therapeutic approach used. For example, cell-surface proteins, such as receptors, are

5   desirable targets for cytotoxic antibodies that can be generated against the polypeptides disclosed herein.

The sequence information disclosed herein, as derived from the GenBank submissions, can readily be utilized by those skilled in the art to

10   prepare the corresponding full-length polypeptide by peptide synthesis. The same is true for either the polynucleotides or polypeptides disclosed herein for use in the methods of the invention.

The present invention relates to an isolated polypeptide, encoded by

15   one of the polynucleotide transcripts disclosed herein, comprising an amino acid sequence homologous to an amino acid sequence selected from the group consisting of SEQ ID NO: 19 - 26, wherein any difference between amino acid sequence in the isolated polypeptide and the sequence of SEQ ID NO: 19 - 26 is due solely to conservative amino acid substitutions and

20   wherein said isolated polypeptide comprises at least one immunogenic fragment. In a preferred embodiment, the present invention encompasses an isolated polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO: 19 - 26.

25   Methods of producing recombinant cells and vectors useful in preparing the polynucleotides and polypeptides disclosed herein are well known to those skilled in the molecular biology art. See, for example, Sambrook, et al., Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor, N.Y., (1989), Wu et al., Methods in Gene Biotechnology (CRC

30   Press, New York, NY, 1997), and Recombinant Gene Expression Protocols, in Methods in Molecular Biology, Vol. 62, (Tuan, ed., Humana Press, Totowa, NJ, 1997), the disclosures of which are hereby incorporated by reference.

In one aspect, the present invention relates to a process for identifying an agent that modulates the activity of a cancer-related gene comprising:

(a) contacting a compound with a cell containing a gene that corresponds to a polynucleotide having a sequence selected from the group
5   consisting of SEQ ID NO: 1 - 18 and under conditions promoting the expression of said gene; and

(b) detecting a difference in expression of said gene relative to when said compound is not present

thereby identifying an agent that modulates the activity of a cancer-
10   related gene.

In specific embodiments of such process the cell is a cancer cell and the difference in expression is a decrease in expression. Such polynucleotides may also include those that have sequences identical to SEQ
15   ID NO: 1 - 18.

In another aspect, the present invention relates to a process for identifying an anti-neoplastic agent comprising contacting a cell exhibiting neoplastic activity with a compound first identified as a cancer related gene
20   modulator using an assay process disclosed herein and detecting a decrease in said neoplastic activity after said contacting compared to when said contacting does not occur. Such neoplastic activity may include accelerated cellular replication and/or metastasis, and the decrease in neoplastic activity preferably results from the death of the cell.

25

The present invention also relates to a process for identifying an anti-neoplastic agent comprising administering to an animal exhibiting a cancer condition an effective amount of an agent first identified according to a process of one of one of the assays disclosed according to the invention and
30   detecting a decrease in said cancerous condition.

In specific embodiments of the present invention, the genes useful for the invention comprise genes that correspond to polynucleotides having a sequence selected from SEQ ID NO: 1 - 18, or may comprise the sequence of any of the polynucleotides disclosed herein (where the latter are cDNA

5   sequences).

In accordance with the present invention, such assays rely on methods of determining the activity of the gene in question. Such assays are advantageously based on model cellular systems using cancer cell lines,

10   primary cancer cells, or cancerous tissue samples that are maintained in growth medium and treated with compounds at a single concentration or at a range of concentrations. At specific times after treatment, cellular RNAs are conveniently isolated from the treated cells or tissues, which RNAs are indicative of expression of selected genes. The cellular RNA is then divided

15   and subjected to differential analysis that detects the presence and/or quantity of specific RNA transcripts, which transcripts may then be amplified for detection purposes using standard methodologies, such as, for example, reverse transcriptase polymerase chain reaction (RT-PCR), etc. The presence or absence, or concentration levels, of specific RNA transcripts are

20   determined from these measurements. The polynucleotide sequences disclosed herein are readily used as probes for the detection of such RNA transcripts and thus the measurement of gene activity and expression.

The polynucleotides of the invention can include fully operational genes

25   with attendant control or regulatory sequences or merely a polynucleotide sequence encoding the corresponding polypeptide or an active fragment or analog thereof.

Because expression of the polynucleotide sequences disclosed herein

30   are specific to the cancerous state, useful gene modulation is downward modulation, so that, as a result of exposure to an antineoplastic agent identified by the screening assays herein, the corresponding gene of the

cancerous cell is expressed at a lower level (or not expressed at all) when exposed to the agent as compared to the expression when not exposed to the agent. For example, the gene sequences disclosed herein (SEQ ID NO: 1 - 18) correspond to a gene expressed at a higher level in cells of prostate
5   cancer than in normal prostate cells. Thus, where said chemical agent causes this gene of the tested cell to be expressed at a lower level than the same genes of the reference, this is indicative of downward modulation and indicates that the chemical agent to be tested has anti-neoplastic activity.

10   In carrying out the assays disclosed herein, relative antineoplastic activity may be ascertained by the extent to which a given chemical agent modulates the expression of genes present in a cancerous cell. Thus, a first chemical agent that modulates the expression of a gene associated with the cancerous state (i.e., a gene corresponding to one or more of the polynucleotide transcripts
15   disclosed herein) to a larger degree than a second chemical agent tested by the assays of the invention is thereby deemed to have higher, or more desirable, or more advantageous, anti-neoplastic activity than said second chemical agent.

The gene expression to be measured is commonly assayed using RNA
20   expression as an indicator. Thus, the greater the level of RNA (for example, messenger RNA or mRNA) detected the higher the level of expression of the corresponding gene. Thus, gene expression, either absolute or relative, is determined by the relative expression of the RNAs encoded by such genes.

25   RNA may be isolated from samples in a variety of ways, including lysis and denaturation with a phenolic solution containing a chaotropic agent (e.g., trizol) followed by isopropanol precipitation, ethanol wash, and resuspension in aqueous solution; or lysis and denaturation followed by isolation on solid support, such as a Qiagen resin and reconstitution in aqueous solution; or lysis
30   and denaturation in non-phenolic, aqueous solutions followed by enzymatic conversion of RNA to DNA template copies.

Normally, prior to applying the processes of the invention, steady state RNA expression levels for the genes, and sets of genes, disclosed herein will have been obtained. It is the steady state level of such expression that is affected by potential anti-neoplastic agents as determined herein. Such steady

5   state levels of expression are easily determined by any methods that are sensitive, specific and accurate. Such methods include, but are in no way limited to, real time quantitative polymerase chain reaction (PCR), for example, using a Perkin-Elmer 7700 sequence detection system with gene specific primer probe combinations as designed using any of several commercially available software

10   packages, such as Primer Express software., solid support based hybridization array technology using appropriate internal controls for quantitation, including filter, bead, or microchip based arrays, solid support based hybridization arrays using, for example, chemiluminescent, fluorescent, or electrochemical reaction based detection systems.

15

The gene expression indicative of a cancerous state need not be characteristic of every cell of a given tissue. Thus, the methods disclosed herein are useful for detecting the presence of a cancerous condition within a tissue where less than all cells exhibit the complete pattern. Thus, for

20   example, a selected gene corresponding to the sequence of SEQ ID NO: 1, may be found, using appropriate probes, either DNA or RNA, to be present in as little as 60% of cells derived from a sample of tumorous, or malignant, tissue. In a highly preferred embodiment, such gene pattern is found to be present in at least 100% of cells drawn from a cancerous tissue and absent

25   from at least 100% of a corresponding normal, non-cancerous, tissue sample.

Expression of a gene may be related to copy number, and changes in expression may be measured by determining copy number. Such change in gene copy number may be determined by determining a change in expression

30   of messenger RNA encoded by a particular gene sequence, especially that of SEQ ID NO: 1 - 18. Also in accordance with the present invention, said gene may be a cancer initiating or facilitating gene. In carrying out the methods of

the present invention, a cancer facilitating gene is a gene that, while not directly initiating tumor formation or growth, acts, such as through the actions of its expression product, to direct, enhance, or otherwise facilitate the progress of the cancerous condition, including where such gene acts against

5   genes, or gene expression products, that would otherwise have the effect of decreasing tumor formation and/or growth.

Although the expression of a gene corresponding to a sequence of SEQ ID NO: 1 - 18 may be indicative of a cancerous status for a given cell,

10   the mere presence of such a gene may not alone be sufficient to achieve a malignant condition and thus the level of expression of such gene may also be a significant factor in determining the attainment of a cancerous state. Thus, it becomes essential to also determine the level of expression of a gene as disclosed herein, including substantially similar sequences, as a separate

15   means of diagnosing the presence of a cancerous status for a given cell, groups of cells, or tissues, either in culture or *in situ*.

The level of expression of the polypeptides disclosed herein is also a measure of gene expression, such as polypeptides having sequence identical,

20   or similar to, any polypeptide encoded by a sequence of SEQ ID NO: 1 - 18, especially a polypeptide whose amino acid sequence is the sequence of SEQ ID NO: 19 - 26.

In accordance with the foregoing, the present invention specifically

25   contemplates a method for determining the cancerous status of a cell to be tested, comprising determining the level of expression in said cell of a gene that includes one of the nucleotide sequences selected from the sequences of SEQ ID NO: 1 - 18, including sequences substantially identical to said sequences, or characteristic fragments thereof, or the complements of any of

30   the foregoing and then comparing said expression to that of a cell known to be non-cancerous whereby the difference in said expression indicates that said cell to be tested is cancerous.

In accordance with the invention, although gene expression for a gene that includes as a portion thereof one of the sequences of SEQ ID NO: 1 - 18, is preferably determined by use of a probe that is a fragment of such

5   nucleotide sequence, it is to be understood that the probe may be formed from a different portion of the gene. Expression of the gene may be determined by use of a nucleotide probe that hybridizes to messenger RNA (mRNA) transcribed from a portion of the gene other than the specific nucleotide sequence disclosed herein.

10

It should be noted that there are a variety of different contexts in which genes have been evaluated as being involved in the cancerous process. Thus, some genes may be oncogenes and encode proteins that are directly involved in the cancerous process and thereby promote the occurrence of

15   cancer in an animal. In addition, other genes may serve to suppress the cancerous state in a given cell or cell type and thereby work against a cancerous condition forming in an animal. Other genes may simply be involved either directly or indirectly in the cancerous process or condition and may serve in an ancillary capacity with respect to the cancerous state. All

20   such types of genes are deemed with those to be determined in accordance with the invention as disclosed herein. Thus, the gene determined by said process of the invention may be an oncogene, or the gene determined by said process may be a cancer facilitating gene, the latter including a gene that directly or indirectly affects the cancerous process, either in the promotion of a

25   cancerous condition or in facilitating the progress of cancerous growth or otherwise modulating the growth of cancer cells, either *in vivo* or *ex vivo*. In addition, the gene determined by said process may be a cancer suppressor gene, which gene works either directly or indirectly to suppress the initiation or progress of a cancerous condition. Such genes may work indirectly where

30   their expression alters the activity of some other gene or gene expression product that is itself directly involved in initiating or facilitating the progress of a cancerous condition. For example, a gene that encodes a polypeptide,

either wild or mutant in type, which polypeptide acts to suppress of tumor suppressor gene, or its expression product, will thereby act indirectly to promote tumor growth.

5   As noted previously, polynucleotides encoding the same proteins as any of SEQ ID NO: 1 - 18, regardless of the percent identity of such sequences, are also specifically contemplated by any of the methods of the present invention that rely on any or all of said sequences, regardless of how they are otherwise described or limited. Thus, any such sequences are

10   available for use in carrying out any of the methods disclosed according to the invention. Such sequences also include any open reading frames, as defined herein, present within the sequence of SEQ ID NO: 1 - 18.

Because a gene disclosed according to the invention "corresponds to"

15   a polynucleotide having a sequence of SEQ ID NO: 1 - 18, said gene encodes an RNA (processed or unprocessed, including naturally occurring splice variants and alleles) that is at least 90% identical, preferably at least 95% identical, most preferably at least 98% identical to, and especially identical to, an RNA that would be encoded by, or be complementary to, such as by

20   hybridization with, a polynucleotide having the indicated sequence. In addition, genes including sequences at least 90% identical to a sequence selected from SEQ ID NO: 1 - 18, preferably at least about 95% identical to such a sequence, more preferably at least about 98% identical to such sequence and most preferably comprising such sequence are specifically

25   contemplated by all of the processes of the present invention. Sequences encoding the same proteins as any of these sequences, regardless of the percent identity of such sequences, are also specifically contemplated by any of the methods of the present invention that rely on any or all of said sequences, regardless of how they are otherwise described or limited. The

30   polynucleotide sequences of the invention also include any open reading frames, as defined herein, present within any of the sequences of SEQ ID NO: 1 - 18.

The sequences disclosed herein may be genomic in nature and thus represent the sequence of an actual gene, such as a human gene, or may be a cDNA sequence derived from a messenger RNA (mRNA) and thus

5   represent contiguous exonic sequences derived from a corresponding genomic sequence, or they may be wholly synthetic in origin for purposes of practicing the processes of the invention. Because of the processing that may take place in transforming the initial RNA transcript into the final mRNA, the sequences disclosed herein may represent less than the full genomic

10   sequence. They may also represent sequences derived from ribosomal and transfer RNAs. Consequently, the gene as present in the cell (and representing the genomic sequence) and the polynucleotide transcripts disclosed herein, including cDNA sequences, may be identical or may be such that the cDNAs contain less than the full genomic sequence. Such

15   genes and cDNA sequences are still considered "corresponding sequences" (as defined elsewhere herein) because they both encode the same or related RNA sequences (i.e., related in the sense of being splice variants or RNAs at different stages of processing). Thus, by way of non-limiting example only, a gene that encodes an RNA transcript, which is then processed into a shorter

20   mRNA, is deemed to encode both such RNAs and therefore encodes an RNA complementary to (using the usual Watson-Crick complementarity rules), or that would otherwise be encoded by, a cDNA (for example, a sequence as disclosed herein). Thus, the sequences disclosed herein correspond to genes contained in the cancerous cells (here, prostate cancer) and are used to

25   determine gene activity or expression because they represent the same sequence or are complementary to RNAs encoded by the gene. Such a gene also includes different alleles and splice variants that may occur in the cells used in the methods of the invention, such as where recombinant cells are used to assay for anti-neoplastic agents and such cells have been engineered

30   to express a polynucleotide as disclosed herein, including cells that have been engineered to express such polynucleotides at a higher level than is found in non-engineered cancerous cells or where such recombinant cells

express such polynucleotides only after having been engineered to do so. Such engineering includes genetic engineering, such as where one or more of the polynucleotides disclosed herein has been inserted into the genome of such cell or is present in a vector.

5

Such cells, especially mammalian cells, may also be engineered to express on their surfaces one or more of the polypeptides of the invention for testing with antibodies or other agents capable of masking such polypeptides and thereby removing the cancerous nature of the cell. Such engineering

10   includes both genetic engineering, where the genetic complement of the cells is engineered to express the polypeptide, as well as non-genetic engineering, whereby the cell has been physically manipulated to incorporate a polypeptide of the invention in its plasma membrane, such as by direct insertion using chemical and/or other agents to achieve this result.

15

In accordance with the foregoing, the present invention includes anti-cancer agents that are themselves either polypeptides, or small chemical entities, that affect the cancerous process, including initiation, suppression or facilitation of tumor growth, either *in vivo* or *ex vivo*. Said cancer modulating

20   agent will have the effect of decreasing gene expression.

The present invention thus also relates to a method for treating cancer comprising contacting a cancerous cell with an agent having activity against an expression product encoded by a gene or polynucleotide sequence as

25   disclosed herein, such as one having, or corresponding to, the nucleotide sequence of SEQ ID NO: 1 - 18. The present invention also relates to a process for treating cancer comprising contacting a cancerous cell with an agent having activity against an expression product encoded by a gene or polynucleotide sequence corresponding to a sequence selected from the

30   group consisting of SEQ ID NO: 1 - 18. In one such embodiment, the cancerous cell is contacted *in vivo*. In another such embodiment, said agent has affinity for said expression product. In a preferred embodiment, such

agent is an antibody disclosed herein, such as an antibody that is specific or selective for, or otherwise reacts with, a polypeptide of the invention. In a preferred embodiment, the expression product is a polypeptide incorporating an amino acid sequence selected from SEQ ID NO: 19 - 26.

5

The present invention is also directed to such uses of the compositions of polypeptides and antibodies disclosed herein. Such uses include a process for treating cancer in an animal afflicted therewith comprising administering to said animal an amount of an immunogenic composition of one or more of the

10   polypeptides disclosed herein where such amount if an amount sufficient to elicit the production of cytotoxic T lymphocytes specific for a polypeptide of the invention, preferably a polypeptide incorporating a sequence of SEQ ID NO: 19 - 26. In a preferred embodiment, the animal to be so treated is a human patient.

15

The proteins encoded by the genes disclosed herein due to their expression, or elevated expression, in cancer cells, represent highly useful therapeutic targets for "targeted therapies" utilizing such affinity structures as, for example, antibodies coupled to some cytotoxic agent. In such

20   methodology, it is advantageous that nothing need be known about the endogenous ligands or binding partners for such cell surface molecules. Rather, an antibody or equivalent molecule that can specifically recognize the cell surface molecule (which could include an artificial peptide, a surrogate ligand, and the like) that is coupled to some agent that can induce cell death

25   or a block in cell cycling offers therapeutic promise against these proteins. Thus, such approaches include the use of so-called suicide "bullets" against intracellular proteins. For example, monoclonal antibodies may readily by produced by methods well known in the art, for example, the method of Kohler and Milstein (see: Nature, 256:495 (1975).

30

With the advent of methods of molecular biology and recombinant technology, it is now possible to produce antibody molecules by recombinant

means and thereby generate gene sequences that code for specific amino acid sequences found in the polypeptide structure of the antibodies. Such antibodies can be produced by either cloning the gene sequences encoding the polypeptide chains of said antibodies or by direct synthesis of said

5   polypeptide chains, with *in vitro* assembly of the synthesized chains to form active tetrameric (H_2L_2) structures with affinity for specific epitopes and antigenic determinants. This has permitted the ready production of antibodies having sequences characteristic of neutralizing antibodies from different species and sources.

10   Regardless of the source of the antibodies, or how they are recombinantly constructed, or how they are synthesized, *in vitro* or *in vivo*, using transgenic animals, such as cows, goats and sheep, using large cell cultures of laboratory or commercial size, in bioreactors or by direct chemical synthesis employing no living organisms at any stage of the process, all

15   antibodies have a similar overall 3 dimensional structure. This structure is often given as H_2L_2 and refers to the fact that antibodies commonly comprise 2 light (L) amino acid chains and 2 heavy (H) amino acid chains. Both chains have regions capable of interacting with a structurally complementary antigenic target. The regions interacting with the target are referred to as

20   "variable" or "V" regions and are characterized by differences in amino acid sequence from antibodies of different antigenic specificity.

The variable regions of either H or L chains contains the amino acid sequences capable of specifically binding to antigenic targets. Within these

25   sequences are smaller sequences dubbed "hypervariable" because of their extreme variability between antibodies of differing specificity. Such hypervariable regions are also referred to as "complementarity determining regions" or "CDR" regions. These CDR regions account for the basic specificity of the antibody for a particular antigenic determinant structure.

30

The CDRs represent non-contiguous stretches of amino acids within the variable regions but, regardless of species, the positional locations of

these critical amino acid sequences within the variable heavy and light chain regions have been found to have similar locations within the amino acid sequences of the variable chains. The variable heavy and light chains of all antibodies each have 3 CDR regions, each non-contiguous with the others

5   (termed L1, L2, L3, H1, H2, H3) for the respective light (L) and heavy (H) chains. The accepted CDR regions have been described by Kabat et al., *J. Biol. Chem.* **252**:6609-6616 (1977).

       In all mammalian species, antibody polypeptides contain constant (i.e.,

10   highly conserved) and variable regions, and, within the latter, there are the CDRs and the so-called "framework regions" made up of amino acid sequences within the variable region of the heavy or light chain but outside the CDRs.

15   The antibodies disclosed according to the invention may also be wholly synthetic, wherein the polypeptide chains of the antibodies are synthesized and, possibly, optimized for binding to the polypeptides disclosed herein as being receptors. Such antibodies may be chimeric or humanized antibodies and may be fully tetrameric in structure, or may be dimeric and comprise only

20   a single heavy and a single light chain. Such antibodies may also include fragments, such as Fab and $F(ab_2)'$ fragments, capable of reacting with and binding to any of the polypeptides disclosed herein as being receptors.

       In one aspect, the present invention relates to immunoglobulins, or

25   antibodies, as described herein, that react with, especially where they are specific for, the polypeptides having amino acid sequences as disclosed herein, preferably those having an amino acid sequence of one of SEQ ID NO: 19 - 26. Such antibodies may commonly be in the form of a composition, especially a pharmaceutical composition. Such antibodies, by themselves,

30   may have therapeutic value in that they are able to bind to, and thereby tie up, surface sites on cancerous cells. Where such sites have some type of function to perform (i.e., where they are surface enzymes, or channel

structures, or structures that otherwise facilitate, actively or passively, the transport of nutrients and other vital materials to the cell. Such nutrients serve to facilitate the growth and replication of the cell and molecules that bind to such sites and thereby interfere with such activities can prove to have a

5   therapeutic effect in that the result of such binding is to remove sources of nutrients from such cells, thereby interfering with growth and replication. In like manner, such binding may serve to remove vital enzyme activities from the cell's functional repertoire, thereby also interfering with viability and/or the ability of the cell to multiply or metastasize. In addition, by binding to such

10   surface sites, the antibodies may serve to prevent the cells from reacting to environmental agents, such as cytokines and the like, that may facilitate growth, replication and metastasis, thereby further reducing the cancerous status of such cell and ameliorating the cancerous condition in a patient, even without proving fatal to the cell or cells so affected.

15   The methods of the present invention also include processes wherein the cancer cell is contacted *in vivo* as well as *ex vivo* with an agent that comprises a portion, or is part of an overall molecular structure, having affinity for an expression product of a gene corresponding to a polynucleotide sequence as disclosed herein, preferably where the expression product is a

20   cell surface structure, most preferably a polypeptide as disclosed herein, such as one that comprises an amino acid sequence of SEQ ID NO: 19 - 26. In one such embodiment, said portion having affinity for said expression product is an antibody, especially where said expression product is a polypeptide or oligopeptide or comprises an oligopeptide portion, or comprises a polypeptide.

25

In another aspect, the present invention also relates to an antibody that reacts with a polypeptide as disclosed herein, preferably a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO: 19 - 26. Such an antibody may be polyclonal, monoclonal,

30   recombinant or synthetic in origin. In one such embodiment, said antibody is associated, either covalently or non-covalently, with a cytotoxic agent, for example, an apoptotic agent. It is thus contemplated that the antibody acts a

targeted vector for guiding an associated therapeutic agent to a cancerous cell, such as a cell expressing a polypeptide homologous to, if not identical to, a polypeptide as disclosed herein.

5   Where the cytotoxic agent is itself a polypeptide, said may be linked directly to an antibody specific for a surface target on a cancer cell, such as where the polypeptide represents an extension of the amino acid chain of the antibody. In alternative embodiments, such molecules may be covalently linked through a linker sequence of long or short duration, such as an amino

10   acid sequence of 5 to 10 residues in length. Where the cytotoxic agents is some small organic molecule, such as a small organic compound, or some type of apoptotic agent, this may be covalently bonded to the antibody molecule or may be attached by some other type of non-covalent linkage, including hydrophobic and electrostatic linkages. Methods for forming such

15   linkages, especially covalent linkages, are well known to those skilled in the art.

        The antibodies disclosed herein may also serve as targeting vectors for much larger structures, such as liposomes. In one such embodiment, an

20   antibody is part of, or otherwise linked to, or associated with, a membranous structure, preferably a liposome or possibly some type of cellular organelle, which acts as a reservoir for a cytotoxic agent, such as ricin. The antibody then acts to target said liposome to a cancerous tissue in an animal, whereupon the liposome provides a source of cytotoxic agents for localized

25   treatment of a solid tumor or other type of neoplasm.

        The present invention further encompasses an immunogenic composition comprising a polypeptide disclosed herein, as well as compositions formed using antibodies specific for these polypeptides.

30

        Methods well known in the art for making formulations are found in, for example, Remington: The Science and Practice of Pharmacy, (19th ed.) Ed.

A.R. Gennaro, 1995, Mack Publishing Company, Easton, PA. Formulations for parenteral administration may, for example, contain excipients, sterile water, or saline, polyalkylene glycols such as polyethylene glycol, oils of vegetable origin, or hydrogenated napthalenes. Biocompatible, biodegradable

5   lactide polymer, lactide/glycolide copolymer, or polyoxyethylene-polyoxypropylene copolymers may be used to control the release of the compounds. Other potentially useful parenteral delivery systems for agonists of the invention include ethylenevinyl acetate copolymer particles, osmotic pumps, implantable infusion systems, and liposomes. Formulations for

10   inhalation may contain excipients, or example, lactose, or may be aqueous solutions containing, for example, polyoxyethylene-9-lauryl ether, glycocholate and deoxycholate, or may be oily solutions for administration in the form of nasal drops, or as a gel. It should be noted that, where the therapeutic agent to be administered is an immunoconjugate, these sometimes contain

15   chemical linkages that are somewhat labile in aqueous media and therefor must be stored prior to administration is a more stable environment, such as in the form of a lyophilized powder.

        Such an agent can be a single molecular structure, comprising both

20   affinity portion and anti-cancer activity portions, wherein said portions are derived from separate molecules, or molecular structures, possessing such activity when separated and wherein such agent has been formed by combining said portions into one larger molecular structure, such as where said portions are combined into the form of an adduct. Said anti-cancer and

25   affinity portions may be joined covalently, such as in the form of a single polypeptide, or polypeptide-like, structure or may be joined non-covalently, such as by hydrophobic or electrostatic interactions, such structures having been formed by means well known in the chemical arts. Alternatively, the anti-cancer and affinity portions may be formed from separate domains of a single

30   molecule that exhibits, as part of the same chemical structure, more than one activity wherein one of the activities is against cancer cells, or tumor formation

or growth, and the other activity is affinity for an expression product produced by expression of genes related to the cancerous process or condition.

In one embodiment of the present invention, a chemical agent, such as

5   a protein or other polypeptide, is joined to an agent, such as an antibody, having affinity for an expression product of a cancerous cell, such as a polypeptide or protein encoded by a gene related to the cancerous process, preferably a gene as disclosed herein according to the present invention, most preferably a polypeptide sequence disclosed herein. Thus, where the

10   presence of said expression product is essential to tumor initiation and/or growth, binding of said agent to said expression product will have the effect of negating said tumor promoting activity. In one such embodiment, said agent is an apoptosis-inducing agent that induces cell suicide, thereby killing the cancer cell and halting tumor growth.

15   Other genes within the cancer cell that are regulated in a manner similar to that of the genes disclosed herein and thus change their expression in a coordinated way in response to chemical compounds represent genes that are located within a common metabolic, signaling, physiological, or functional pathway so that by analyzing and identifying such commonly

20   regulated groups of genes (groups that include the gene, or similar sequences, disclosed according to the invention, one can (a) assign known genes and novel genes to specific pathways and (b) identify specific functions and functional roles for novel genes that are grouped into pathways with genes for which their functions are already characterized or described. For

25   example, one might identify a group of 10 genes, at least one of which is the gene as disclosed herein, that change expression in a coordinated fashion and for which the function of one, such as the polypeptide encoded by the sequence disclosed herein, is known then the other genes are thereby implicated in a similar function or pathway and may thus play a role in the

30   cancer-initiating or cancer-facilitating process. In the same way, if a gene were found in normal cells but not in cancer cells, or happens to be expressed at a higher level in normal as opposed to cancer cells, then a similar

conclusion may be drawn as to its involvement in cancer, or other diseases. Therefore, the processes disclosed according to the present invention at once provide a novel means of assigning function to genes, i.e. a novel method of functional genomics, and a means for identifying chemical compounds that

5   have potential therapeutic effects on specific cellular pathways. Such chemical compounds may have therapeutic relevance to a variety of diseases outside of cancer as well, in cases where such diseases are known or are demonstrated to involve the specific cellular pathway that is affected.

10   The polypeptides disclosed herein, preferably those of SEQ ID NO: 19 - 26, also find use as vaccines in that, where the polypeptide represents a surface protein present on a cancer cell, such polypeptide may be administered to an animal, especially a human being, for purposes of activating cytotoxic T lymphocytes (CTLs) that will be specific for, and act to

15   lyze, cancer cells in said animal. Where used as vaccines, such polypeptides are present in the form of a pharmaceutical composition. The present invention may also employ polypeptides that have the same, or similar, immunogenic character as the polypeptides of SEQ ID NO: 19 - 26 and thereby elicit the same, or similar, immunogenic response after administration

20   to an animal, such as an animal at risk of developing cancer, or afflicted therewith. Thus, the polypeptides disclosed according to the invention will commonly find use as immunogenic compositions.

        Expression of a gene corresponding to a polynucleotide disclosed

25   herein, when in normal tissues, may indicate a predisposition towards development of prostate cancer. The encoded polypeptide might then present a potentially useful cell surface target for therapeutic molecules such as cytolytic antibodies, or antibodies attached to cytotoxic, or cytolytic, agents. .

30   The present invention specifically contemplates use of antibodies against the polypeptides encoded by the polynucleotides corresponding to the genes disclosed herein, whereby said antibodies are conjugates to one or

more cytotoxic agents so that the antibodies serve to target the conjugated immunotoxins to a region of cancerous activity, such as a solid tumor. For many known cytotoxic agents, lack of selectivity has presented a drawback to their use as therapeutic agents in the treatment of malignancies. For example,

5   the class of two-chain toxins, consisting of a binding subunit (or B-chain) linked to a toxic subunit (A-chain) are extremely cytotoxic. Thus, such agents as ricin, a protein isolated from castor beans, kills cells at very low concentrations (even less than 10^{-11} M) by inactivating ribosomes in said cells (see, for example, Lord et al., Ricin: structure, mode of action, and some

10   current applications. *Faseb J*, **8**: 201-208 (1994), and Blättler et al., Realizing the full potential of immunotoxins. *Cancer Cells*, **1**: 50-55 (1989)). While isolated A-chains of protein toxins that functionally resemble ricin A-chain are only weakly cytotoxic for intact cells (in the concentration range of 10^{-7} to 10^{-6} M), they are very potent cytotoxic agents inside the cells. Thus, a single

15   molecule of the A-subunit of diphtheria toxin can kill a cell once inside (see: Yamaizumi et al., One molecule of diphtheria toxin fragment A introduced into a cell can kill the cell. Cell, *15*: 245-250, 1978).

The present invention solves this selectivity problem by using antibodies specific for antigens present on cancer cells to target the cytotoxins

20   to said cells. In addition, use of antibodies decreases toxicity because the antibodies are non-toxic until they reach the tumor and, because the cytotoxin is bound to the antibody, it is presented with less opportunity to cause damage to non-targeted tissues.

25   In addition, use of such antibodies alone can provide therapeutic effects on the tumor through the antibody-dependent cellular cytotoxic response (ADCC) and complement-mediated cell lysis mechanisms.

A number of recombinant immunotoxins (for example, consisting of Fv

30   regions of cancer specific antibodies fused to truncated bacterial toxins) are well known (see, for example, Smyth et al., Specific targeting of chlorambucil to tumors with the use of monoclonal antibodies, *J. Natl. Cancer Inst.*,

76(3):503-510 (1986); Cho et al., Single-chain Fv/folate conjugates mediate efficient lysis of folate-receptor-positive tumor cells, *Bioconjug. Chem.*, 8(3):338-346 (1997)). As noted in the literature, these may contain, for example, a truncated version of *Pseudomonas* exotoxin as a toxic moiety but

5   the toxin is modified in such a manner that by itself it does not bind to normal human cells, but it retains all other functions of cytotoxicity. Here, recombinant antibody fragments target the modified toxin to cancer cells which are killed, such as by direct inhibition of protein synthesis, or by concomitant induction of apoptosis. Cells that are not recognized by the antibody fragment, because

10 they do not carry the cancer antigen, are not affected. Good activity and specificity has been observed for many recombinant immunotoxins in *in vitro* assays using cultured cancer cells as well as in animal tumor models. Ongoing clinical trials provide examples where the promising pre-clinical data correlate with successful results in experimental cancer therapy. (see, for

15 example, Brinkmann U., Recombinant antibody fragments and immunotoxin fusions for cancer therapy, *In Vivo* (2000) **14**:21-27).

While the safety of employing immunoconjugates in humans has been established, *in vivo* therapeutic results have been less impressive. Because

20 clinical use of mouse MAbs in humans is limited by the development of a foreign anti-globulin immune response by the human host, genetically engineered chimeric human-mouse MAbs have been developed by replacing the mouse Fc region with the human constant region. In other cases, the mouse antibodies have been "humanized" by replacing the framework regions

25 of variable domains of rodent antibodies by their human equivalents. Such humanized and engineered antibodies can even be structurally arranged to have specificities and effector functions determined by design and which characteristics do not appear in nature. The development of bispecific antibodies, having different binding ends so that more than one antigenic site

30 can be bound, have proven useful in targeting cancer cells. Thus, such antibody specificity has been improved by chemical coupling to various agents such as bacterial or plant toxins, radionuclides or cytotoxic drugs and other agents. (see, for example, Bodey, B. et al). Genetically engineered

monoclonal antibodies for direct anti-neoplastic treatment and cancer cell specific delivery of chemotherapeutic agents. _Curr Pharm Des_ (2000) Feb;6(3):261-76). See also, Garnett, M. C., Targeted drug conjugates: principles and progress. _Adv. Drug Deliv. Rev._ (2001 Dec 17) 53(2):171-216;

5   Brinkmann et al., Recombinant immunotoxins for cancer therapy. _Expert Opin Biol Ther._ (2001) 1(4):693-702.

Among the cytotoxic agents specifically contemplated for use as immunoconjugates according to the present invention are Calicheamicin, a

10   highly toxic enediyne antibiotic isolated from _Micromonospora echinospora ssp. Calichensis_, and which binds to the minor groove of DNA to induce double strand breaks and cell death (see: Lee et al., Calicheamicins, a novel family of antitumor antibiotics. 1. Chemistry and partial structure of calichemicin g_1. _J Am Chem Soc_, **109**: 3464-3466 (1987);

15   Zein et al., Calicheamicin gamma 1I: an antitumor antibiotic that cleaves double-stranded DNA site specifically, _Science_, **240**: 1198-1201 (1988)). Useful derivatives of the calicheamicins include mylotarg and 138H11-Camθ. Mylotarg is an immunoconjugate of a humanized anti-CD33 antibody (CD33 being found in leukemic cells of most patients with acute myeloid leukemia)

20   and N-acetyl gamma colicheamicin dimethyl hydrazide, the latter of which is readily coupled to an antibody of the present invention (in place of the anti-CD33 but which can also be humanized by substitution of human framework regions into the antibody during production as described elsewhere herein) to form an immunoconjugate of the invention. (see: Hamann et al. Gemtuzumab

25   Ozogamicin, A Potent and Selective Anti-CD33 Antibody-Calicheamicin Conjugate for Treatment of Acute Myeloid Leukemia, _Bioconjug. Chem._ **13**, 47-58 (2002)) For use with 138H11-Camθ, 138H11 is an anti-γ-glutamyl transferase antibody coupled to theta calicheamicin through a disulfide linkage and found useful _in vitro_ against cultured renal cell carcinoma cells.

30   (see: Knoll et al., Targeted therapy of experimental renal cell carcinoma with a novel conjugate of monoclonal antibody 138H11 and calicheamicin θ_1^1, _Cancer Res_, **60**: 6089-6094 (2000) The same linkage may be utilized to link

this cytotoxic agent to an antibody of the present invention, thereby forming a targeting structure for prostate cancer cells.

Also useful in forming the immunoconjugates of the invention is DC1, a
5   disulfide-containing analog of adozelesin, that kills cells by binding to the
minor groove of DNA, followed by alkylation of adenine bases. Adozelesin is a
structural analog of CC-1065, an anti-tumor antibiotic isolated from microbial
fermentation of *Streptomyces zelensis*, and is about 1,000 fold more toxic to
cultured cell lines that other DNA interacting agents, such as cis-platin and
10   doxorubicin. This agent is readily linked to antibodies through the disulfide
bond of adozelesin. (see: Chari et al., Enhancement of the selectivity and
antitumor efficacy of a CC-1065 analogue through immunoconjugate
formation, *Cancer Res*, **55**: 4079-4084 (1995)).

Maytansine, a highly cytotoxic microtubular inhibitor isolated from the
15   shrub *Maytenus serrata* found to have little value in human clinical trials, is
much more effective in its derivatized form, denoted DM1, containing a
disulfide bond to facilitate linkage to antibodies, is up to 10-fold more cytotoxic
(see: Chari et al., Immunoconjugates containing novel maytansinoids:
promising anticancer drugs, *Cancer Res*, **52**: 127-131 (1992)). These same *in*
20   *vitro* studies showed that up to four DM1 molecules could be linked to a single
immunoglobulin without destroying the binding affinity. Such conjugates have
been used against breast cancer antigens, such as the *neu/HER2/erbB-2*
antigen. (see: Goldmacher et al., Immunogen, Inc., (2002) *in press*); also see
Liu, C. et al., Eradication of large colon tumor xenografts by targeted delivery
25   of maytansinoids, *Proc. Natl. Acad. Sci. USA*, **93**, 8618-8623 (1996)). For
example, Liu et al. (1996) describes formation of an immunoconjugate of the
maytansinoid cytotoxin DM1 and C242 antibody, a murine IgG1
immunoglobulin, available from Pharmacia and which has affinity for a mucin-
like glycoprotein variably expressed by human colorectal cancers. The latter
30   immunoconjugate was prepared according to Chari et al., *Cancer Res.*,
52:127-131 (1992) and was found to be highly cytotoxic against cultured colon
cancer cells as well as showing anti-tumor effects *in vivo* in mice bearing

subcutaneous COLO 205 human colon tumor xenografts using doses well below the maximum tolerated dose.

5   In addition, there are a variety of protein toxins (cytotoxic proteins), which include a number of different classes, such as those that inhibit protein synthesis: ribosome-inactivating proteins of plant origin, such as ricin, abrin, gelonin, and a number of others, and bacterial toxins such as pseudomonas exotoxin and diphtheria toxin.

10   Another useful class is the one including taxol, taxotere, and taxoids. Specific examples include paclitaxel (taxol), its analog docetaxel (taxotere), and derivatives thereof. The first two are clinical drugs used in treating a number of tumors while the taxoids act to induce cell death by inhibiting the de-polymerization of tubulin. Such agents are readily linked to antibodies
15   through disulfide bonds without disadvantageous effects on binding specificity.

       In one instance, a truncated Pseudomonas exotoxin was fused to an anti-CD22 variable fragment and used successfully to treat patients with
20   chemotherapy-resistant hairy-cell leukemia. (see: Kreitman et al., Efficacy of the anti-CD22 recombinant immunotoxin BL22 in chemotherapy-resistant hairy-cell leukemia, N Engl J Med, **345**: 241-247 (2001)) Conversely, the cancer-linked peptides of the present invention offer the opportunity to prepare antibodies, recombinant or otherwise, against the appropriate
25   antigens to target solid tumors, preferably those of malignancies of prostate tissue, using the same or similar cytotoxic conjugates. Thus, many of the previously used immunoconjugates have been formed using antibodies against general antigenic sites linked to cancers whereas the antibodies formed using the peptides disclosed herein are more specific and target the
30   antibody-cytotoxic agent to a particular tissue or organ, thus further reducing toxicity and other undesirable side effects.

In addition, the immunoconjugates formed using the antibodies prepared against the cancer-linked antigens disclosed herein can be formed by any type of chemical coupling. Thus, the cytotoxic agent of choice, along with the immunoglobulin, can be coupled by any type of chemical linkage,

5   covalent or non-covalent, including electrostatic linkage, to form the immunoconjugates of the present invention.

When used as immunoconjugates, the antitumor agents of the present invention represent a class of pro-drugs that are relatively non-toxic when first

10   administered to an animal (due mostly to the stability of the immunoconjugate), such as a human patient, but which are targeted by the conjugated immunoglobulin to a cancer cell where they then exhibit good toxicity. The tumor-related, associated, or linked, antigens, preferably those presented herein, serve as targets for the antibodies (monoclonal,

15   recombinant, and the like) specific for said antigens. The end result is the release of active cytotoxic agent inside the cell after binding of the immunoglobulin portion of the immunoconjugate.

The cited references describe a number of useful procedures for the

20   chemical linkage of cytotoxic agents to immunoglobulins and the disclosures of all such references cited herein are hereby incorporated by reference in their entirety. For other reviews see Ghetie et al., Immunotoxins in the therapy of cancer: from bench to clinic, Pharmacol Ther, 63: 209-234 (1994), Pietersz et al. The use of monoclonal antibody immunoconjugates in cancer therapy,

25   Adv Exp Med Biol, 353:169-179 (1994), and Pietersz, G. A. The linkage of cytotoxic drugs to monoclonal antibodies for the treatment of cancer, Bioconjug Chem, 1:89-95 (1990).

Thus, the present invention provides highly useful cancer-associated

30   antigens for generation of antibodies for linkage to a number of different cytotoxic agents which are already known to have some in vitro toxicity and possess chemical groups available for linkage to antibodies.

Because the polypeptides disclosed herein are diagnostic for prostate cancer, the present invention offers a method for diagnosing prostate cancer in a patient comprising determining the presence of a cancer-related

5   polypeptide in one or more prostate cells of said patient wherein said cancer-related polypeptide comprises a polypeptide that reacts with an antibody specific for a polypeptide having an amino acid sequence selected from SEQ ID NO: 19 – 26. In a preferred embodiment thereof, the cancer-related polypeptide comprises a polypeptide of SEQ ID NO: 19 – 26. In a highly

10 preferred embodiment, the cancer-related polypeptide is a member selected from the group consisting of SEQ ID NO: 19 – 26.

The present invention also relates to a process that comprises a method for producing a product, including the generation of test data,

15 comprising identifying an agent according to one of the disclosed processes for identifying such an agent (i.e., the therapeutic agents identified according to the assay procedures disclosed herein) wherein said product is the data collected with respect to said agent as a result of said identification process, or assay, and wherein said data is sufficient to convey the chemical character

20 and/or structure and/or properties of said agent. For example, the present invention specifically contemplates a situation whereby a user of an assay of the invention may use the assay to screen for compounds having the desired enzyme modulating activity and, having identified the compound, then conveys that information (i.e., information as to structure, dosage, etc) to

25 another user who then utilizes the information to reproduce the agent and administer it for therapeutic or research purposes according to the invention. For example, the user of the assay (user 1) may screen a number of test compounds without knowing the structure or identity of the compounds (such as where a number of code numbers are used the first user is simply given samples labeled with said code numbers) and, after performing the screening

30 samples labeled with said code numbers) and, after performing the screening process, using one or more assay processes of the present invention, then imparts to a second user (user 2), verbally or in writing or some equivalent

fashion, sufficient information to identify the compounds having a particular modulating activity (for example, the code number with the corresponding results). This transmission of information from user 1 to user 2 is specifically contemplated by the present invention.

5

It should be cautioned that, in carrying out the procedures of the present invention as disclosed herein, whether to form immunoconjugates or screen for other antitumor agents using the genes and polypeptides disclosed herein, any reference to particular buffers, media, reagents, cells, culture

10 conditions and the like are not intended to be limiting, but are to be read so as to include all related materials that one of ordinary skill in the art would recognize as being of interest or value in the particular context in which that discussion is presented. For example, it is often possible to substitute one buffer system or culture medium for another and still achieve similar, if not

15 identical, results. Those of skill in the art will have sufficient knowledge of such systems and methodologies so as to be able, without undue experimentation, to make such substitutions as will optimally serve their purposes in using the methods and procedures disclosed herein.

20 The present invention will now be further described by way of the following non-limiting example. In applying the disclosure of the example, it should be kept clearly in mind that other and different embodiments of the methods disclosed according to the present invention will no doubt suggest themselves to those of skill in the relevant art. The following example shows

25 how a potential anti-neoplastic agent may be identified using one or more of the genes disclosed herein.

EXAMPLE

30

Determination of Gene Inhibitory Activity of an Anti-neoplastic Agent

38

SW480 cells are grown to a density of 10^5 cells/cm^2 in Leibovitz's L-15 medium supplemented with 2 mM L-glutamine (90%) and 10% fetal bovine serum. The cells are collected after treatment with 0.25% trypsin, 0.02% EDTA at 37°C for 2 to 5 minutes. The trypsinized cells are then diluted with 30

5   ml growth medium and plated at a density of 50,000 cells per well in a 96 well plate (100 µl/well). The following day, cells are treated with either compound buffer alone, or compound buffer containing a chemical agent to be tested, for 24 hours. The media is then removed, the cells lysed and the RNA recovered using the RNAeasy reagents and protocol obtained from Qiagen. RNA is

10   quantitated and 10 ng of sample in 1 µl are added to 24 µl of Taqman reaction mix containing 1X PCR buffer, RNAsin, reverse transcriptase, nucleoside triphosphates, amplitaq gold, tween 20, glycerol, bovine serum albumin (BSA) and specific PCR primers and probes for a reference gene (18S RNA) and a test gene (Gene X). Reverse transcription is then carried out at 48°C for 30

15   minutes. The sample is then applied to a Perlin Elmer 7700 sequence detector and heat denatured for 10 minutes at 95°C. Amplification is performed through 40 cycles using 15 seconds annealing at 60°C followed by a 60 second extension at 72°C and 30 second denaturation at 95°C. Data files are then captured and the data analyzed with the appropriate baseline

20   windows and thresholds.

The quantitative difference between the target and reference gene is then calculated and a relative expression value determined for all of the samples used. In this way, the ability of a chemotherapeutic agent to

25   effectively and selectively reduce the activity of a cancer-specific gene is readily ascertained. The overall expression of the cancer-specific gene, as modulated by one chemical agent relative to another, is also determined. Chemical agents having the most effect in reducing gene activity are thereby identified as the most anti-neoplastic.

30

WO 03/104404

References:

Walter A. Blättler and Ravi Chari: Drugs to enhance the therapeutic potency of anti-cancer antibodies: antibody-drug conjugates as tumor-activated
5   prodrugs. In *Anticancer Agents - Frontiers in Cancer Chemotherapy* (Iwao Ojima, Gregory D. Vite, Karl-Heinz Altmann, Eds.), American Chemical Society, pp. 317-338 (2001).

Dan L. Longo, Patricia L. Duffey, John G. Gribben, Elaine S. Jaffe, Brendan D. Curti, Barry L. Gause, John E. Janik, Virginia M. Braman, Dixie Esseltine,
10   Wyndham H. Wilson, Dwight Kaufman, Robert E. Wittes, Lee M. Nadler, and Walter J. Urba: Combination chemotherapy followed by an Immunotoxin (Anti-B4-blocked Ricin) in patients with indolent lymphoma: results of a Phase II study. *Cancer J. 6*, 146-150 (2000).

Walter A. Blättler and John M. Lambert: Preclinical immunotoxin development.
15   In *Monoclonal Antibody-Based Therapy of Cancer* (M. Grossbard, Ed.), Marcel Dekker, Inc. NY, NY, pp. 1-22 (1998).

Ravi V. J. Chari: Targeted delivery of chemotherapeutics: tumor-activated prodrug therapy. In *Advanced Drug Delivery Reviews*, Elsevier Science B.V., pp. 89-104 (1998).

20   David T. Scadden, David P. Schenkein, Zale Bernstein, Barry Luskey, John Doweiko, Anil Tulpule, and Alexandra M. Levine: Immunotoxin combined with chemotherapy for patients with AIDS-related Non-Hodgkin's Lymphoma. *Cancer 83*, 2580-2587 (1998).

Changnian Liu and Ravi VJ Chari: The development of antibody delivery
25   systems to target cancer with highly potent maytansinoids. *Exp. Opi. Invest. Drugs 6*, 169-172 (1997).

A. C. Goulet, Viktor S. Goldmacher, John M. Lambert, C. Baron, Dennis C. Roy and E. Kouassi: Conjugation of blocked ricin to an anti-CD19 monoclonal antibody increases antibody-induced cell calcium mobilization and CD19
30   internalization. *Blood 90*, 2364-2375 (1997).

Changnian Liu, John M. Lambert, Beverly A. Teicher, Walter A. Blättler, and Rosemary O'Connor: Cure of multidrug-resistant human B-cell lymphoma xenografts by combinations of anti-B4-blocked ricin and chemotherapeutic drugs. *Blood 87*, 3892-3898 (1996).

35   Rajeeva Singh, Lana Kats, Walter A. Blättler, and John M. Lambert: Formation of N-Substituted 2-Iminothiolanes when amino groups in proteins and peptides are modified by 2-Iminothiolane. *Anal. Biochem. 236*, 114-125 (1996).

Changnian Liu, B. Mitra Tadayoni, Lizabeth A. Bourret, Kristin M. Mattocks,
40   Susan M. Derr, Wayne C. Widdison, Nancy L. Kedersha, Pamela D. Ariniello, Victor S. Goldmacher, John M. Lambert, Walter A. Blättler, and Ravi V.J.

Chari: Eradication of large colon tumor xenografts by targeted delivery of maytansinoids. *Proc. Natl. Acad. Sci. USA 93*, 8618-8623 (1996).

Denis C. Roy, Sophie Ouellet, Christiane Le Houiller, Pamela D. Ariniello, Claude Perreault and John M. Lambert: Elimination of neuroblastoma and small-cell lung cancer cells with an anti-neural cell adhesion molecule immunotoxin. *J. Natl. Cancer Inst. 88*, 1136-1145 (1996).

Walter A. Blättler, Ravi V.J. Chari and John M. Lambert: Immunoconjugates. In Cancer Therapeutics: Experimental and Clinical Agents. (B. Teicher, Ed.), Humana Press, Totowa, NJ, pp. 371-394 (1996).

Michael L Grossbard, John M. Lambert, Victor S. Goldmacher, Arnold S. Freedman, Jeanne Kinsella, Danny P. Ducello, Susan N. Rabinowe, Laura Elisea, Felice Carol, James A. Taylor, Walter A. Blättler, Carol L. Epstein, and Lee M. Nadler: Anti-B4-blocked Ricin: A phase I trial of 7 day continuous infusion in patients with B-cell neoplasms. *J. Clin. Oncol. 11*, 726-737 (1993).

Michael L. Grossbard, John G. Gribben, Arnold S. Freedman, John M. Lambert, Jeanne Kinsella, Susan N. Rabinowe, Laura Eliseo, James A. Taylor, Walter A. Blättler, Carol L. Epstein, and Lee M. Nadler: Adjuvant immunotoxin therapy with anti-B4-blocked ricin following autologous bone marrow transplantation for patients with B-cell Non-Hodgkin's lymphoma. *Blood 81*, 2263-2271 (1993).

Sudhir A. Shah, Patricia M. Halloran, Cynthia A. Ferris, Beth A. Levine, Lizabeth A. Bourret, Victor S. Goldmacher, and Walter A. Blättler: Anti-B4-blocked Ricin immunotoxin shows therapeutic efficacy in four different SCID mouse tumor models. *Cancer Res. 53*, 1360-1367 (1993).

Ravi V.J. Chari, Bridget A. Martell, Jonathan L. Gross, Sherilyn B. Cook, Sudhir A. Shah, Walter A. Blättler, Sara J. McKenzie, and Victor S. Goldmacher: Immunoconjugates containing novel maytansinoids: promising anti-cancer drugs. *Cancer Res. 52*, 127-131 (1992).

John M. Lambert, Peter D. Senter, Annie Yau-Young, Walter A. Blättler, and Victor S. Goldmacher: Purified immunotoxins that are reactive with human lymphoid cells. *J. Biol. Chem. 250*, 12035-12041 (1985).

WHAT IS CLAIMED IS:

1. A process for identifying an agent that modulates the activity of a cancer-related gene comprising:

5   (a) contacting a compound with a cell containing a gene that corresponds to a polynucleotide having a sequence selected from the group consisting of SEQ ID NO: 1 - 18 and under conditions promoting the expression of said gene; and

(b) detecting a difference in expression of said gene relative to when
10   said compound is not present

thereby identifying an agent that modulates the activity of a cancer-related gene.

2 The process of claim 1 wherein said gene has a sequence selected
15   from the group consisting of SEQ ID NO: 1 - 18.

3. The process of claim 1 wherein the cell is a cancer cell and the difference in expression is a decrease in expression.

20   4. The process of claim 3 wherein said cancer cell is a prostate cancer cell.

5. A process for identifying an anti-neoplastic agent comprising contacting a cell exhibiting neoplastic activity with a compound first identified
25   as a cancer related gene modulator using the process of claim 1 and detecting a decrease in said neoplastic activity after said contacting compared to when said contacting does not occur.

6. The process of claim 5 wherein said neoplastic activity is
30   accelerated cellular replication.

7. The process of claim 5 wherein said decrease in neoplastic activity results from the death of the cell.

8. A process for identifying an anti-neoplastic agent comprising administering to an animal exhibiting a cancer condition an effective amount of an agent first identified according to the process of claim 1 and detecting a decrease in said cancerous condition.

9. A process for determining the cancerous status of a cell, comprising determining an increase in the level of expression in said cell of a gene that corresponds to a polynucleotide having a sequence selected from the group consisting of SEQ ID NO: 1 - 18 wherein an elevated expression relative to a known non-cancerous cell indicates a cancerous state or potentially cancerous state.

10. The process of claim 9 wherein said elevated expression is due to an increased copy number.

11. An isolated polypeptide comprising an amino acid sequence homologous to an amino acid sequence selected from the group consisting of SEQ ID NO: 19 - 26 wherein any difference between said amino acid sequence and the sequence of SEQ ID NO: 19 - 26 is due solely to conservative amino acid substitutions and wherein said isolated polypeptide comprises at least one immunogenic fragment.

12. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO: 19 - 26.

13. An antibody that reacts with a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO: 19 - 26.

14. The antibody of claim 13 wherein said antibody is a recombinant antibody.

15. The antibody of claim 13 wherein said antibody is a synthetic antibody.

16. The antibody of claim 13 wherein said antibody is a humanized antibody.

17. An immunoconjugate comprising the antibody of claim 13 and a cytotoxic agent.

18. The antibody of claim 17 wherein said cytotoxic agent is a member selected from the group consisting of a calicheamicin, a maytansinoid, an adozelesin, a cytotoxic protein, a taxol, a taxotere, a taxoid and DC1.

19. The immunoconjugate of claim 18 wherein said calicheamicin is calicheamicin γ_1^I, N-acetyl gamma calicheamicin dimethyl hydrazide or calicheamicin θ_1^I.

20. The immunoconjugate of claim 18 wherein said maytansinoid is DM1.

21. The immunoconjugate of claim 18 wherein said cytotoxic protein is ricin, abrin, gelonin, pseudomonas exotoxin or diphtheria toxin.

22. The immunoconjugate of claim 18 wherein said taxol is paclitaxel.

23. The immunoconjugate of claim 18 wherein said taxotere is docetaxel.

24. A process for treating cancer comprising contacting a cancerous cell *in vivo* with an agent having activity against an expression product encoded by a gene sequence selected from the group consisting of SEQ ID NO: 1 - 18.

5

25. The process of claim 24 wherein said agent is an antibody of claim 13.

26. The process of claim 24 wherein said agent is an immunoconjugate

10   of claim 17.

27. An immunogenic composition comprising a polypeptide of claim 11.

28. An immunogenic composition comprising a polypeptide of claim 12.

15

29. The process of claim 24 wherein said cancer is prostate cancer.

30. A process for treating cancer in an animal afflicted therewith comprising administering to said animal an amount of an immunogenic

20   composition of claim 27 sufficient to elicit the production of cytotoxic T lymphocytes specific for the polypeptide of claim 11.

31. A process for treating cancer in an animal afflicted therewith comprising administering to said animal an amount of an immunogenic

25   composition of claim 28 sufficient to elicit the production of cytotoxic T lymphocytes specific for the polypeptide of claim 12.

32. A process for treating a cancerous condition in an animal afflicted therewith comprising administering to said animal a therapeutically effective

30   amount of an agent first identified as having anti-neoplastic activity using the process of claim 8.

33. A process for protecting an animal against cancer comprising administering to an animal at risk of developing cancer a therapeutically effective amount of an agent first identified as having anti-neoplastic activity using the process of claim 8.

5

34. The process of claim 30 wherein said animal is a human being.

35. The process of claim 30 wherein said cancer is prostate cancer.

10   36. A method for diagnosing prostate cancer in a patient comprising determining the presence of a cancer-related polypeptide in one or more prostate cells of said patient wherein said cancer-related polypeptide comprises a polypeptide that reacts with an antibody specific for a polypeptide having an amino acid sequence selected from SEQ ID NO: 19 – 26.

15

37. The method of claim 36 wherein said cancer-related polypeptide comprises a polypeptide of SEQ ID NO: 19 – 26.

38. The method of claim 36 wherein said cancer-related polypeptide is
20   a member selected from the group consisting of SEQ ID NO: 19 – 26.

39. A method for producing test data with respect to the gene modulating activity of a compound comprising:

(a) contacting a compound with a cell containing a polynucleotide
25   comprising a nucleotide sequence corresponding to a gene whose expression is increased in a cancerous cell over that in a non-cancerous cell and under conditions wherein said polynucleotide is being expressed,

(b) determining a change in expression of polynucleotides as a result of said contacting, and

30   (c) producing test data with respect to the gene modulating activity of said compound based on a decrease in the expression of the determined

gene whose expression is otherwise increased in a cancerous cell over that in a non-cancerous cell indicating gene modulating activity.

SEQUENCE LISTING

<110> Avalon Pharmaceuticals

<120> Cancer-Linked Gene as Target for Chemotherapy

<130> 689290-166

<140>
<141>

<150> US/60/386,651
<151> 2002-06-06

<160> 26

<170> PatentIn version 3.0

<210> 1
<211> 5271
<212> DNA
<213> Homo sapiens

<400> 1
ggacaggcgt ggcggccgga gccccagcat ccctgcttga ggtccaggag cggagcccgc   60
ggccaccgcc gcctgatcag cgcgacccog gcccgcgccc gccccgcccg gcaagatgct   120
gcccgtgtac caggaggtga agcccaaccc gctgcaggac gcgaacatct gctcacgcgt   180
gttcttctgg tggctcaatc ccttgtttaa aattggccat aaacggagat tagaggaaga   240
tgatatgtat tcagtgctgc cagaagaccg ctcacagcac cttggagagg agttgcaagg   300
gttctgggat aaagaagttt taagagctga gaatgacgca cagaagcctt ctttaacaag   360
agcaatcata aagtgttact ggaaatctta tttagttttg ggaattttta cgttaattga   420
ggaaagtgcc aaagtaatcc agcccatatt tttgggaaaa attattaatt attttgaaaa   480
ttatgatccc atggattctg tggctttgaa cacagcgtac gcctatgcca cggtgctgac   540
tttttgcacg ctcattttgg ctatactgca tcacttatat ttttatcacg ttcagtgtgc   600
tgggatgagg ttacgagtag ccatgtgcca tatgatttat cggaaggcac ttcgtcttag   660
taacatggcc atggggaaga caaccacagg ccagatagtc aatctgctgt ccaatgatgt   720
gaacaagttt gatcaggtga cagtgttctt acacttcctg tgggcaggac cactgcaggc   780
gatcgcagtg actgccctac tctggatgga gataggaata tcgtgccttg ctgggatggc   840
agttctaatc attctcctgc ccttgcaaag ctgtttgggg aagttgttct catcactgag   900
gagtaaaact gcaactttca cggatgccag gatcaggacc atgaatgaag ttataactgg   960
tataaggata ataaaaatgt acgcctggga aaagtcattt tcaaatctta ttaccaattt   1020
gagaaagaag gagatttcca agattctgag aagttcctgc ctcaggggga tgaatttggc   1080
ttcgtttttc agtgcaagca aaatcatcgt gtttgtgacc ttcaccacct acgtgctcct   1140
cggcagtgtg atcacagcca gccgcgtgtt cgtggcagtg acgctgtatg gggctgtgcg   1200
gctgacggtt accctcttct tcccctcagc cattgagagg gtgtcagagg caatcgtcag   1260
catccgaaga atccagacct ttttgctact tgatgagata tcacagcgca accgtcagct   1320
gccgtcagat ggtaaaaaga tggtgcatgt gcaggatttt actgcttttt gggataaggc   1380
atcagagacc ccaactctac aaggcctttc ctttactgtc agacctggcg aattgttagc   1440
tgtggtcggc cccgtgggag cagggaagtc atcactgtta agtgccgtgc tcgggggaatt   1500
ggccccaagt cacgggctgg tcagcgtgca tggaagaatt gcctatgtgt ctcagcagcc   1560
ctgggtgttc tcggaactc tgaggagtaa tattttattt gggaagaaat atgaaaagga   1620
acgatatgaa aaagtcataa aggcttgtgc tctgaaaaag gatttacagc tgttggagga   1680
tggtgatctg actgtgatag agatcggggg aaccacgctg agtggagggc agaaagcacg   1740
ggtaaacctt gcaagagcag tgtatcaaga tgctgacatc tatctcctgg acgatcctct   1800
cagtgcagta gatgcggaag ttagcagaca cttgttcgaa ctgtgtattt gtcaaatttt   1860
gcatgagaag atcacaattt tagtgactca tcagttgcag tacctcaaag ctgcaagtca   1920
gattctgata ttgaaagatg gtaaaatggt gcagaagggg acttacactg agttcctaaa   1980
atctggtata gattttggct ccctttttaaa gaaggataat gaggaaagtg aacaacctcc   2040

                                     1

```
agttccagga actcccacac taaggaatcg taccttctca gagtcttcgg tttggtctca    2100
acaatcttct agaccctcct tgaaagatgg tgctctggag agccaagata cagagaatgt    2160
cccagttaca ctatcagagg agaaccgttc tgaaggaaaa gttggttttc aggcctataa    2220
gaattacttc agagctggtg ctcactggat tgtcttcatt ttccttattc tcctaaacac    2280
tgcagctcag gttgcctatg tgcttcaaga ttggtggctt tcatactggg caaacaaaca    2340
aagtatgcta aatgtcactg taaatggagg aggaaatgta accgagaagc tagatcttaa    2400
ctggtactta ggaatttatt caggtttaac tgtagctacc gttctttttg gcatagcaag    2460
atctctattg gtattctacg tccttgttaa ctcttcacaa actttgcaca caaaatgtt    2520
tgagtcaatt ctgaaagctc cggtattatt ctttgataga aatccaatag gaagaatttt    2580
aaatcgtttc tccaaagaca ttggacactt ggatgatttg ctgccgctga cgtttttaga    2640
tttcatccag acattgctac aagtggttgg tgtggtctct gtggctgtgg ccgtgattcc    2700
ttggatcgca ataccccttgg ttcccctgg aatcattttc attttttcttc ggcgatattt    2760
tttggaaacg tcaagagatg tgaagcgcct ggaatctaca actcggagtc cagtgttttc    2820
ccacttgtca tcttctctcc aggggctctg gaccatccgg gcatacaaag cagaagagag    2880
gtgtcaggaa ctgtttgatg cacaccagga tttacattca gaggcttggt tcttgttttt    2940
gacaacgtcc cgctggttcg ccgtccgtct ggatgccatc tgtgccatgt ttgtcatcat    3000
cgttgccttt gggtccctga ttctggcaaa aactctggat gccgggcagg ttggtttggc    3060
actgtcctat gccctcacgc tcatggggat gtttcagtgg tgtgttcgac aaagtgctga    3120
agttgagaat atgatgatct cagtagaaag ggtcattgaa tacacagacc ttgaaaaaga    3180
agcaccttgg gaatatcaga aacgcccacc accagcctgg ccccatgaag gagtgataat    3240
ctttgacaat gtgaacttca tgtacagtcc aggtgggcct ctggtactga agcatctgac    3300
agcactcatt aaatcacaag aaaaggttgg cattgtggga agaaccggag ctggaaaaag    3360
ttccctcatc tcagccccttt ttagattgtc agaacccgaa ggtaaaattt ggattgataa    3420
gatcttgaca actgaaattg gacttcacga tttaaggaag aaaatgtcaa tcatacctca    3480
ggaacctgtt ttgttcactg gaacaatgag gaaaaacctg gatcccttta atgagcacac    3540
ggatgaggaa ctgtggaatg ccttacaaga ggtacaactt aaagaaacca ttgaagatct    3600
tcctggtaaa atggatactg aattagcaga atcaggatcc aattttagtg ttggacaaag    3660
acaactggtg tgccttgcca gggcaattct caggaaaaat cagatattga ttattgatga    3720
agcgacggca aatgtggatc caagaactga tgagttaata caaaaaaaaa tccgggagaa    3780
atttgcccac tgcaccgtgc taaccattgc acacagattg aacaccatta ttgacagcga    3840
caagataatg gtttttagatt caggaagact gaaagaatat gatgagccgt atgtttttgct    3900
gcaaaataaa gagagcctat tttacaagat ggtgcaacaa ctgggcaagg cagaagccgc    3960
tgccctcact gaaacagcaa aacaggtata cttcaaaaga aattatccac atattggtca    4020
cactgaccac atggttacaa acacttccaa tggacagccc tcgaccttaa ctattttcga    4080
gacagcactg tgaatccaac caaaatgtca agtccgttcc gaaggcattt tccactagtt    4140
tttggactat gtaaaccaca ttgtactttt ttttactttg gcaacaaata tttatacata    4200
caagatgcta gttcatttga atatttctcc caacttatcc aaggatctcc agctctaaca    4260
aaatggttta tttttatta aatgtcaata gttgtttttt aaaatccaaa tcagaggtgc    4320
aggccaccag ttaaatgccg tctatcaggt tttgtgcctt aagagactac agagtcaaag    4380
ctcatttta aaggagtagg acaaagttgt cacaggtttt tgttgttgtt tttattgccc    4440
ccaaaattac atgttaattt ccatttatat cagggattct atttacttga agactgtgaa    4500
gttgccattt tgtctcattg ttttctttga cataactagg atccattatt tcccctgaag    4560
gcttcttgtt agaaaatagt acagttacaa ccaataggaa caacaaaaag aaaaagtttg    4620
tgacattgta gtagggagtg tgtacccctt actccccatc aaaaaaaaaa atggatacat    4680
ggttaaagga tagaagggca atatttttatc atatgttcta aaagagaagg aagagaaaat    4740
actactttct caaaatggaa gcccttaaag gtgctttgat actgaaggac acaaatgtga    4800
ccgtccatcc tcctttagag ttgcatgact tggacacggt aactgttgca gttttagact    4860
cagcattgtg acacttccca agaaggccaa acctctaacc gacattcctg aaatacgtgg    4920
cattattctt ttttggattt ctcatttatg gaaggctaac cctctgttga ccgtaagcct    4980
tttggtttgg gctgtattga aatcctttct aaattgcatg aataggctct gctaacgtga    5040
tgagacaaac tgaaaattat tgcaagcatt gactataatt atgcagtacg ttctcaggat    5100
gcatccaggg gttcattttc atgagcctgt ccaggttagt ttactcctga ccactaatag    5160
cattgtcatt tgggctttct gttgaatgaa tcaacaaacc acaatacttc ctgggacctt    5220
ttgtactttta tttgaactat gagtctgtat taactgcttg gggtaccgag c          5271
```

<210>   2
<211>   5284
<212>   DNA

2

<213> Homo sapiens

<400> 2

```
ggacaggcgt ggcggccgga gccccagcat ccctgcttga ggtccaggag cggagcccgc    60
ggccaccgcc gcctgatcag cgcgaccccg gcccgcgccc gccccgcccg gcaagatgct   120
gcccgtgtac caggaggtga agcccaaccc gctgcaggac gcgaacatct gctcacgcgt   180
gttcttctgg tggctcaatc ccttgtttaa aattggccat aaacggagat tagaggaaga   240
tgatatgtat tcagtgctgc cagaagaccg ctcacagcac cttggagagg agttgcaagg   300
gttctgggat aaagaagttt taagagctga gaatgacgca cagaagcctt ctttaacaag   360
agcaatcata aagtgttact ggaaatctta tttagttttg ggaattttta cgttaattga   420
ggaaagtgcc aaagtaatcc agcccatatt tttgggaaaa attattaatt attttgaaaa   480
ttatgatccc atggattctg tggctttgaa cacagcgtac gcctatgcca cggtgctgac   540
ttttttgcacg ctcatttttgg ctatactgca tcacttatat ttttatcacg ttcagtgtgc   600
tgggatgagg ttacgagtag ccatgtgcca tatgatttat cggaaggcac ttcgtcttag   660
taacatggcc atggggaaga caaccacagg ccagatagtc aatctgctgt ccaatgatgt   720
gaacaagttt gatcaggtga cagtgttctt acacttcctg tgggcaggac cactgcaggc   780
gatcgcagtg actgccctac tctggatgga gataggaata tcgtgccttg ctgggatggc   840
agttctaatc attctcctgc ccttgcaaag ctgttttggg aagttgttct catcactgag   900
gagtaaaaact gcaactttca cggatgccag gatcaggacc atgaatgaag ttataactgg   960
tataaggata ataaaaatgt acgcctggga aaagtcattt tcaaatctta ttaccaattt  1020
gagaaagaag gagatttcca agattctgag aagttcctgc ctcaggggga tgaatttggc  1080
ttcgtttttc agtgcaagca aaatcatcgt gtttgtgacc ttcaccacct acgtgctcct  1140
cggcagtgtg atcacagcca gccgcgtgtt cgtggcagtg acgctgtatg gggctgtgcg  1200
gctgacggtt accctcttct tcccctcagc cattgagagg gtgtcagagg caatcgtcag  1260
catccgaaga atccagacct ttttgctact tgatgagata tcacagcgca accgtcagct  1320
gccgtcagat ggtaaaaaga tggtgcatgt gcaggatttt actgcttttt gggataaggc  1380
atcagagacc ccaactctac aaggcctttc ctttactgtc agacctggcg aattgttagc  1440
tgtggtcggc cccgtgggag cagggaagtc atcactgtta agtgccgtgc tcggggaatt  1500
ggccccaagt cacgggctgg tcagcgtgca tggaagaatt gcctatgtgt ctcagcagcc  1560
ctgggtgttc tcgggaactc tgaggagtaa tattttattt gggaagaaat atgaaaagga  1620
acgatatgaa aaagtcataa aggcttgtgc tctgaaaaag gatttacagc tgttggagga  1680
tggtgatctg actgtgatag gagatcgggg aaccacgctg agtggagggc agaaagcacg  1740
ggtaaacctt gcaagagcag tgtatcaaga tgctgacatc tatctcctgg acgatcctct  1800
cagtgcagta gatgcggaag ttagcagaca cttgttcgaa ctgtgtattt gtcaaatttt  1860
gcatgagaag atcacaattt tagtgactca tcagttgcag tacctcaaag ctgcaagtca  1920
gattctgata ttgaaagatg gtaaaatggt gcagaagggg acttacactg agttcctaaa  1980
atctggtata gattttggct ccctttttaaa gaaggataat gaggaaagtg aacaacctcc  2040
agttccagga actcccacac taaggaatcg taccttctca gagtcttcgg tttggtctca  2100
acaatcttct agaccctcct tgaaagatgg tgctctggag agccaagata cagagaatgt  2160
cccagttaca ctatcagagg agaaccgttc tgaaggaaaa gttggttttc aggcctataa  2220
gaattacttc agagctggtg ctcactggat tgtcttcatt ttccttattc tcctaaacac  2280
tgcagctcag gttgcctatg tgcttcaaga ttggtggctt tcatactggg caaacaaaca  2340
aagtatgcta aatgtcactg taaatggagg aggaaatgta accgagaagc tagatcttaa  2400
ctggtactta ggaatttatt caggtttaac tgtagctacc gttctttttg gcatagcaag  2460
atctctattg gtattctacg tccttgttaa ctcttcacaa acttttgcaca acaaaatgtt  2520
tgagtcaatt ctgaaagctc cggtattatt ctttgataga aatccaatag aagaattttt  2580
aaatcgtttc tccaaagaca ttggacactt ggatgatttg ctgccgctga cgttttttaga  2640
tttcatccag acattgctac aagtggttgg tgtggtctct gtggctgtgg ccgtgattcc  2700
ttggatcgca atacccttgg ttcccccttgg aatcatttttc attttttcttc ggcgatattt  2760
tttggaaacg tcaagagatg tgaagcgcct ggaatctaca actcggagtc cagtgttttc  2820
ccacttgtca tcttctctcc aggggctctg gaccatccgg gcatacaaag cagaagagag  2880
gtgtcaggaa ctgtttgatg cacaccagga tttacattca gaggcttggt tcttgttttt  2940
gacaacgtcc cgctggttcg ccgtccgtct ggatgccatc tgtgccatgt ttgtcatcat  3000
cgttgccttt gggtccctga ttctggcaaa aactctggat gccgggcagg ttggtttggc  3060
actgtcctat gccctcacgc tcatgggggat gtttcagtgg tgtgttcgac aaagtgctga  3120
agttgagaat atgatgatct cagtagaaag ggtcattgaa tacacagacc ttgaaaaaga  3180
agcaccttgg gaatatcaga aacgcccacc accagcctgg cccatgaag gagtgataat  3240
ctttgacaat gtgaacttca gtacagtcc aggtgggcct ctggtactga agcatctgac  3300
agcactcatt aaatcacaag aaaaggttgg cattgtggga agaaccggag ctggaaaaag  3360
```

```
ttccctcatc tcagcccttt ttagattgtc agaacccgaa ggtaaaattt ggattgataa  3420
gatcttgaca actgaaattg gacttcacga tttaaggaag aaaatgtcaa tcatacctca  3480
ggaacctgtt ttgttcactg gaacaatgag gaaaaacctg gatcccttta atgagcacac  3540
ggatgaggaa ctgtggaatg ccttacaaga ggtacaactt aaagaaacca ttgaagatct  3600
tcctggtaaa atggatactg aattagcaga atcaggatcc aattttagtg ttggacaaag  3660
acaactggtg tgccttgcca gggcaattct caggaaaaat cagatattga ttattgatga  3720
agcgacggca aatgtggatc caagaactga tgagttaata caaaaaaaaa tccgggagaa  3780
atttgcccac tgcaccgtgc taaccattgc acacagattg aacaccatta ttgacagcga  3840
caagataatg gttttagatt caggaagact gaaagaatat gatgagccgt atgtttttgct  3900
gcaaaataaa gagagcctat tttacaagat ggtgcaacaa ctgggcaagg cagaagccgc  3960
tgccctcact gaaacagcaa aacaggtata cttcaaaaga aattatccac atattggtca  4020
cactgaccac atggttacaa acacttccaa tggacagccc tcgaccttaa ctattttcga  4080
gacagcactg tgaatccaac caaaatgtca agtccgttcc gaaggcattt tccactagtt  4140
tttggactat gtaaaccaca ttgtactttt ttttactttg gcaacaaata tttatacata  4200
caagatgcta gttcatttga atatttctcc caacttatcc aaggatctcc agctctaaca  4260
aaatggttta tttttatta aatgtcaata gttgttttttt aaaatccaaa tcagaggtgc  4320
aggccaccag ttaaatgccg tctatcaggt tttgtgcctt aagagactac agagtcaaag  4380
ctcattttta aaggagtagg acaaagttgt cacaggtttt tgttgttgtt tttattgccc  4440
ccaaaattac atgttaattt ccatttatat cagggattct atttacttga agactgtgaa  4500
gttgccattt tgtctcattg ttttctttga cataactagg atccattatt tccctgaag  4560
gcttcttgtt agaaaatagt acagttacaa ccaataggaa caacaaaaag aaaaagtttg  4620
tgacattgta gtagggagtg tgtacccctt actccccatc aaaaaaaaaa atggatacat  4680
ggttaaagga tagaagggca atattttatc atatgttcta aaagagaagg aagagaaat  4740
actactttct caaaatggaa gcccttaaag gtgctttgat actgaaggac acaaatgtga  4800
ccgtccatcc tcctttagag ttgcatgact tggacacggt aactgttgca gttttagact  4860
cagcattgtg acacttccca agaaggccaa acctctaacc gacattcctg aaatacgtgg  4920
cattattctt ttttggattt ctcatttatg gaaggctaac cctctgttga ccgtaagcct  4980
tttggtttgg gctgtattga aatcctttct aaattgcatg aataggctct tggcttaaac  5040
ccggttgaat gagacaaact gagacaactt taaatttggc ccaaaggcca aactttgtcg  5100
ggaacgccct ttttccgtgc acaacaccga cgacccccac ccttttctac aaacccaaac  5160
aacgtgtaca accttttcgggg ccccaacgca cagcccctcc caccgcgccg cccacgaccg  5220
cactaacgcc ccggccactt ctacaactcc cctctcccccc acaagggggcg gacaacttgg  5280
ggca                                                               5284
```

```
<210>  3
<211>  3521
<212>  DNA
<213>  Homo sapiens

<400>  3
ggacaggcgt ggcggccgga gccccagcat ccctgcttga ggtccaggag cggagcccgc   60
ggccaccgcc gcctgatcag cgcgaccccg gcccgcgccc gccccgcccg gcaagatgct  120
gcccgtgtac caggaggtga gcccaaccc gctgcaggac gcgaacatct gctcacgcgt  180
gttcttctgg tggctcaatc ccttgtttaa aattggccat aaacggagat tagaggaaga  240
tgatatgtat tcagtgctgc cagaagaccg ctcacagcac cttggagagg agttgcaagg  300
gttctgggat aaagaagttt taagagctga gaatgacgca cagaagcctt ctttaacaag  360
agcaatcata aagtgttact ggaaatctta tttagttttg ggaattttta cgttaattga  420
ggaaagtgcc aaagtaatcc agcccatatt tttgggaaaa attattaatt attttgaaaa  480
ttatgatccc atggattctg tggctttgaa cacagcgtac gcctatgcca cggtgctgac  540
tttttgcacg ctcatttttgg ctatactgca tcacttatat ttttatcacg ttcagtgtgc  600
tgggatgagg ttacgagtag ccatgtgcca tatgatttat cggaaggcac ttcgtcttag  660
taacatggcc atggggaaga caaccacagg ccagatagtc aatctgctgt ccaatgatgt  720
gaacaagttt gatcaggtga cagtgttctt acacttcctg tgggcaggac cactgcaggc  780
gatcgcagtg actgccctac tctggatgga gataggaata tcgtgccttg ctgggatggc  840
agttctaatc attctcctgc ccttgcaaag ctgttttggg aagttgttct catcactgag  900
gagtaaaact gcaactttca cggatgccag gatcaggacc atgaatgaag ttataactgg  960
tataaggata ataaaaatgt acgcctggga aaagtcattt tcaaatctta ttaccaattt 1020
gagaaagaag gagatttcca agattctgag aagttcctgc ctcaggggga tgaatttggc 1080
```

                                      4

```
ttcgtttttc agtgcaagca aaatcatcgt gtttgtgacc ttcaccacct acgtgctcct   1140
cggcagtgtg atcacagcca gccgcgtgtt cgtggcagtg acgctgtatg gggctgtgcg   1200
gctgacggtt accctcttct tcccctcagc cattgagagg gtgtcagagg caatcgtcag   1260
catccgaaga atccagacct ttttgctact tgatgagata tcacagcgca accgtcagct   1320
gccgtcagat ggtaaaaaga tggtgcatgt gcaggatttt actgcttttt gggataaggc   1380
atcagagacc ccaactctac aaggcctttc ctttactgtc agacctggcg aattgttagc   1440
tgtggtcggc cccgtgggag cagggaagtc atcactgtta agtgccgtgc tcggggaatt   1500
ggccccaagt cacgggctgg tcagcgtgca tggaagaatt gcctatgtgt ctcagcagcc·  1560
ctgggtgttc tcgggaactc tgaggagtaa tattttattt gggaagaaat atgaaaagga   1620
acgatatgaa aaagtcataa aggcttgtgc tctgaaaaag gatttacagc tgttggagga   1680
tggtgatctg actgtgatag agatcggggg aaccacgctg agtggagggc agaaagcacg   1740
ggtaaacctt gcaagagcag tgtatcaaga tgctgacatc tatctcctgg acgatcctct   1800
cagtgcagta gatgcggaag ttagcagaca cttgttcgaa ctgtgtattt gtcaaatttt   1860
gcatgagaag atcacaattt tagtgactca tcagttgcag tacctcaaag ctgcaagtca   1920
gattctgata ttgaaagatg gtaaaatggt gcagaagggg acttacactg agttcctaaa   1980
atctggtata gattttggct ccctttttaaa gaaggataat gaggaaagtg aacaacctcc   2040
agttccagga actcccacac taaggaatcg taccttctca gagtcttcgg tttggtctca   2100
acaatcttct agaccctcct tgaaagatgg tgctctggag agccaagata cagagaatgt   2160
cccagttaca ctatcagagg agaaccgttc tgaaggaaaa gttggttttc aggcctataa   2220
gaattacttc agagctggtg ctcactggat tgtcttcatt ttccttattc tcctaaacac   2280
tgcagctcag gttgcctatg tgcttcaaga ttggtggctt tcatactggg caaacaaaca   2340
aagtatgcta aatgtcactg taaatggagg aggaaatgta accgagaagc tagatcttaa   2400
ctggtactta ggaatttatt caggtttaac tgtagctacc gttctttttg gcatagcaag   2460
atctctattg gtattctacg tccttgttaa ctcttcacaa actttgcaca acaaaatgtt   2520
tgagtcaatt ctgaaagctc cggtattatt ctttgataga aatccaatag gaagaatttt   2580
aaatcgtttc tccaaagaca ttggacactt ggatgatttg ctgccgctga cgttttttaga   2640
tttcatccag acattgctac aagtggttgg tgtggtctct gtggctgtgg ccgtgattcc   2700
ttggatcgca ataccccttgg ttccccttgg aatcatttttc attttttcttc ggcgatattt   2760
tttggaaacg tcaagagatg tgaagcgcct ggaatctaca actcggagtc cagtgttttc   2820
ccacttgtca tcttctctcc aggggctctg gaccatccgg gcatacaaag cagaagagag   2880
gtgtcaggaa ctgtttgatg cacaccagga tttacattca ggtctgtcta tttctggaaa   2940
tggtttcaaa ggacaggatc tgctgctttt tctggcttaa cctccctggt cagatgagct   3000
tgttggcagg cacctcaggg cccagcccca catcctgctt tttttgcgca tcttcgtctc   3060
cttccccatc cccttcttct gggcttggac tcctctcgga atcgctgctg cacgttccct   3120
ctttgtgccc cctgactggc ttgtgggggtc tttctctcac catgcatggc ctgtggctct   3180
gtcagcaaac gttttctgca aaaggccacc tagtaagtgt tttaggcatt gtgtgccatt   3240
attgtctctg ttgtgattaa tcacctttgc cattgtagca caaaagcaga tgtcaataca   3300
ttggcaaatg gctgtgacta tgtttcaata aaactttatt tacaaaaata ggtgctggct   3360
cacgtgctgt agtttgctga ccccttgcat cagtgaagat tttttcaata tggaaatttc   3420
agaatttcca ccaaatctac aaaattcaga tccgtcattc caatctttat tttatcttgc   3480
tgattaaatg gcactttaaa aaccacatgc taaaaggaaa a                       3521
```

```
<210>   4
<211>   3284
<212>   DNA
<213>   Homo sapiens

<400>   4
caagcagggc atgagttcac caccccagta acaacagtgc attacaatac ttaataatgg   60
aacaaggaga gtagtatttt agtctcaaaa gcattggcac ttacaaccta ttgctttgtt   120
tttagtggtt ccagctgtgt gttgcattgc acgtagaaag tggaataatg taatgggctt   180
tgaaaccata ataatgaatg tctgaataat gacattattt cttgcgtttg taatactgtt   240
aattaaatct atgtcgatcc tgttggaatt cataaaatca tctaaaaatt tttctaaata   300
tacagtgttg ttttccccat tgtatcttga tctcaagcaa caaatggtaa aagtatagct   360
attaatgtca ttaaatgtga attgtttcaa cattatgaag ggttcctctt ggtaagtggc   420
agaaggagcc aggcttaggt ttgaagtgag actgacttta ttccttcttc cctccttacg   480
tagctctgga tgccgggcag gttggtttgg cactgtccta tgccctcacg ctcatgggga   540
tgtttcagtg gtgtgttcga caaagtgctg aagttgagaa tatgatgatc tcagtagaaa   600
```

```
gggtcattga atacacagac cttgaaaaag aagcaccttg ggaatatcag aaacgcccac    660
caccagcctg gccccatgaa ggagtgataa tctttgacaa tgtgaacttc atgtacagtc    720
caggtgggcc tctggtactg aagcatctga cagcactcat taaatcacaa gaaaaggttg    780
gcattgtggg aagaaccgga gctggaaaaa gttccctcat ctcagccctt tttagattgt    840
cagaacccga aggtaaaatt tggattgata agatcttgac aactgaaatt ggacttcacg    900
atttaaggaa gaaaatgtca atcatacctc aggaacctgt tttgttcact ggaacaatga    960
ggaaaaacct ggatcccttt aatgagcaca cggatgagga actgtggaat gccttacaag   1020
aggtacaact taaagaaacc attgaagatc ttcctggtaa aatggatact gaattagcag   1080
aatcaggatc caattttagt gttggacaaa gacaactggt gtgccttgcc agggcaattc   1140
tcaggaaaaa tcagatattg attattgatg aagcgacggc aaatgtggat ccaagaactg   1200
atgagttaat acaaaaaaaa atccgggaga aatttgccca ctgcaccgtg ctaaccattg   1260
cacacagatt gaacaccatt attgacagcg acaagataat ggttttagat tcaggaagac   1320
tgaaagaata tgatgagccg tatgttttgc tgcaaaataa agagagccta ttttacaaga   1380
tggtgcaaca actgggcaag gcagaagccg ctgccctcac tgaaacagca aaacaggtat   1440
acttcaaaag aaattatcca catattggtc acactgacca catggttaca aacacttcca   1500
atggacagcc ctcgacctta actattttcg agacagcact gtgaatccaa ccaaaatgtc   1560
aagtccgttc cgaaggcatt ttccactagt ttttggacta tgtaaaccac attgtacttt   1620
ttttttacttt ggcaacaaat atttatacat acaagatgct agttcatttg aatatttctc   1680
ccaacttatc caaggatctc cagctctaac aaaatggttt atttttattt aaatgtcaat   1740
agttgttttt taaaatccaa atcagaggtg caggccacca gttaaatgcc gtctatcagg   1800
ttttgtgcct taagagacta cagagtcaaa gctcattttt aaaggagtag gacaaagttg   1860
tcacaggttt ttgttgttgt ttttattgcc cccaaaatta catgttaatt tccatttata   1920
tcaggattc tatttacttg aagactgtga agttgccatt ttgtctcatt gttttctttg   1980
acataactag gatccattat ttcccctgaa ggcttcttgt tagaaaatag tacagttaca   2040
accaatagga acaacaaaaa gaaaaagttt gtgacattgt agtagggagt gtgtacccct   2100
tactccccat caaaaaaaaa aatggataca tggttaaagg atagaagggc aatattttat   2160
catatgttct aaaagagaag gaagagaaaa tactactttc tcaaaatgga agcccttaaa   2220
ggtgctttga tactgaagga cacaaatgtg accgtccatc ctcctttaga gttgcatgac   2280
ttggacacgg taactgttgc agtttttagac tcagcattgt gacacttccc aagaaggcca   2340
aacctctaac cgacattcct gaaatacgtg gcattattct tttttggatt tctcatttat   2400
ggaaggctaa ccctctgttg accgtaagcc ttttggtttg ggctgtattg aaatcctttc   2460
taaattgcat gaataggctc tgctaacgtg atgagacaaa ctgaaaatta ttgcaagcat   2520
tgactataat tatgcagtac gttctcagga tgcatccagg ggttcatttt catgagcctg   2580
tccaggttag tttactcctg accactaata gcattgtcat ttgggctttc tgttgaatga   2640
atcaacaaac cacaatactt cctgggacct tttgtacttt atttgaacta tgagtcttta   2700
attttttcctg atgatggtgg ctgtaatatg ttgagttcag tttactaaag gttttactat   2760
tatggtttga agtggagtct catgacctct cagaataagg tgtcacctcc ctgaaattgc   2820
atatatgtat atagacatgc acacgtgtgc atttgtttgt atacatatat ttgtccttcg   2880
tatagcaagt tttttgctca tcagcagaga gcaacagatg ttttattgag tgaagcctta   2940
aaaagcacac accacacaca gctaactgcc aaaatacatt gaccgtagta gctgttcaac   3000
tcctagtact tagaaataca cgtatggtta atgttcagtc caacaaacca cacacagtaa   3060
atgtttatta atagtcatgg ttcgtatttt aggtgactga aattgcaaca gtgatcataa   3120
tgaggtttgt taaaacgata gctatattca aaatgtctat atgtttattt ggacttttga   3180
ggttaaagac agtcatataa acgtcctgtt tctgttttaa tgttatcata gaattttta   3240
atgaaactaa attcaattga aataaatgat agttttcatc tcca                    3284
```

<210> 5
<211> 1019
<212> DNA
<213> Homo sapiens

<400> 5
```
ggacaggcgt ggcggccgga gccccagcat ccctgcttga ggtccaggag cggagcccgc     60
ggccaccgcc gcctgatcag cgcgaccccg gcccgcgccc gccccgcccg gcaagatgct    120
gcccgtgtac caggaggtga gcccaacccg ctgcaggac gcgaacatct gctcacgcgt    180
gttcttctgg tggctcaatc ccttgtttaa aattggccat aaacggagat tagaggaaga    240
tgatatgtat tcagtgctgc cagaagaccg ctcacagcac cttggagagg agttgcaagg    300
gttctgggat aaagaagttt taagagctga gaatgacgca cagaagcctt ctttaacaag    360
```

```
agcaatcata aagtgttact ggaaatctta tttagttttg ggaattttta cgttaattga      420
ggaaagtgcc aaagtaatcc agcccatatt tttgggaaaa attattaatt attttgaaaa      480
ttatgatccc atggattctg tggctttgaa cacagcgtac gcctatgcca cggtgctgac      540
tttttgcacg ctcattttgg ctatactgca tcacttatat ttttatcacg ttcagtgtgc      600
tgggatgagg ttacgagtag ccatgtgcca tatgatttat cggaaggcac ttcgtcttag      660
taacatggcc atggggaaga caaccacagg ccagatagtc aatctgctgt ccaatgatgt      720
gaacaagttt gatcaggtga cagtgttctt acacttcctg tgggcaggac cactgcaggc      780
gatcgcagtg actgccctac tctggatgga gataggaata tcgtgccttg ctgggatggc      840
agttctaatc attctcctgc ccttgcaaag ctgttttggg aagttgttct catcactgag      900
gagtaaaact gcaactttca cggatgccag gatcaggacc atgaatgaag ttataactgg      960
tataaggata ataaaaatgt acgcctgtgg taagacgrga tcctyrgagt cactcgcag     1019


<210>   6
<211>   3200
<212>   DNA
<213>   Homo sapiens

<400>   6
ggcagggtcc cctcaggcta ttgggaacat ttattttttcc ttgaaacatg aagcctccta     60
ggaggcctgc gtgtgctgtt tggtggatac gtgaccatgg agcttgggag gccgcaggtt     120
tttctgcacc cagcctcact gagctgtttt ctctgtctgt ttccccagct cggagtccag     180
tgttttccca cttgtcatct tctctccagg ggctctggac catccgggca tacaaagcag     240
aagagaggtg tcaggaactg tttgatgcac accaggattt acattcagag gcttggttct     300
tgtttttgac aacgtcccgc tggttcgccg tccgtctgga tgccatctgt gccatgtttg     360
tcatcatcgt tgcctttggg tccctgattc tggcaaaaac tctggatgcc gggcaggttg     420
gtttggcact gtcctatgcc ctcacgctca tggggatgtt tcagtggtgt gttcgacaaa     480
gtgctgaagt tgagaatatg atgatctcag tagaaagggt cattgaatac acagaccttg     540
aaaaagaagc accttgggaa tatcagaaac gcccaccacc agcctggccc catgaaggag     600
tgataatctt tgacaatgtg aacttcatgt acagtccagg tgggcctctg gtactgaagc     660
atctgacagc actcattaaa tcacaagaaa aggttggcat tgtgggaaga accggagctg     720
gaaaaagttc cctcatctca gccctttttta gattgtcaga acccgaaggt aaaatttgga     780
ttgataagat cttgacaact gaaattggac ttcacgattt aaggaagaaa atgtcaatca     840
tacctcagga acctgttttg ttcactggaa caatgaggaa aaacctggat cccttttaatg     900
agcacacgga tgaggaactg tggaatgcct tacaagaggt acaacttaaa gaaaccattg     960
aagatcttcc tggtaaaatg gatactgaat tagcagaatc aggatccaat tttagtgttg    1020
gacaaagaca actggtgtgc cttgccaggg caattctcag gaaaaatcag atattgatta    1080
ttgatgaagc gacggcaaat gtggatccaa gaactgatga gttaatacaa aaaaaaatcc    1140
gggagaaatt tgcccactgc accgtgctaa ccattgcaca cagattgaac accattattg    1200
acagcgacaa gataatggtt ttagattcag gaagactgaa agaatatgat gagccgtatg    1260
ttttgctgca aaataaagag agcctatttt acaagatggt gcaacaactg ggcaaggcag    1320
aagccgctgc cctcactgaa acagcaaaac aggtatactt caaaagaaat tatccacata    1380
ttggtcacac tgaccacatg gttacaaaca cttccaatgg acagccctcg accttaacta    1440
ttttcgagac agcactgtga atccaaccaa aatgtcaagt ccgttccgaa ggcatttcc     1500
actagttttt ggactatgta aaccacattg tactttttt tactttggca acaaatattt    1560
atacatacaa gatgctagtt catttgaata tttctcccaa cttatccaag gatctccagc    1620
tctaacaaaa tggtttattt ttatttaaat gtcaatagtt gttttttaaa atccaaatca    1680
gaggtgcagg ccaccagtta aatgccgtct atcaggtttt gtgccttaag agactacaga    1740
gtcaaagctc attttaaag gagtaggaca aagttgtcac aggttttttgt tgttgttttt    1800
attgccccca aaattacatg ttaatttcca tttatatcag ggattctatt tacttgaaga    1860
ctgtgaagtt gccattttgt ctcattgttt tctttgacat aactaggatc cattatttcc    1920
cctgaaggct tcttgttaga aaatagtaca gttacaacca ataggaacaa caaaaagaaa    1980
aagtttgtga cattgtagta gggagtgtgt accccttact ccccatcaaa aaaaaaaatg    2040
gatacatggt taaaggatag aagggcaata ttttatcata tgttctaaaa gagaaggaag    2100
agaaaatact actttctcaa aatggaagcc cttaaaggtg ctttgatact gaaggacaca    2160
aatgtgaccg tccatcctcc tttagagttg catgacttgg acacggtaac tgttgcagtt    2220
ttagactcag cattgtgaca cttcccaaga aggccaaacc tctaaccgac attcctgaaa    2280
tacgtggcat tattcttttt tggatttctc atttatggaa ggctaaccct ctgttgaccg    2340
taagcctttt ggtttgggct gtattgaaat cctttctaaa ttgcatgaat aggctctgct    2400
```

7

```
aacgtgatga gacaaactga aaattattgc aagcattgac tataattatg cagtacgttc      2460
tcaggatgca tccaggggtt cattttcatg agcctgtcca ggttagttta ctcctgacca      2520
ctaatagcat tgtcatttgg gctttctgtt gaatgaatca acaaaccaca atacttcctg      2580
ggaccttttg tactttattt gaactatgag tctttaattt ttcctgatga tggtggctgt      2640
aatatgttga gttcagttta ctaaaggttt tactattatg gtttgaagtg gagtctcatg      2700
acctctcaga ataaggtgtc acctccctga aattgcatat atgtatatag acatgcacac      2760
gtgtgcattt gtttgtatac atatatttgt ccttcgtata gcaagttttt tgctcatcag      2820
cagagagcaa cagatgtttt attgagtgaa gccttaaaaa gcacacacca cacacagcta      2880
actgccaaaa tacattgacc gtagtagctg ttcaactcct agtacttaga aatacacgta      2940
tggttaatgt tcagtccaac aaaccacaca cagtaaatgt ttattaatag tcatggttcg      3000
tattttaggt gactgaaatt gcaacagtga tcataatgag gtttgttaaa acgatagcta      3060
tattcaaaat gtctatatgt ttatttggac ttttgaggtt aaagacagtc atataaacgt      3120
cctgtttctg ttttaatgtt atcatagaat tttttaatga aactaaattc aattgaaata      3180
aatgatagtt ttcatctcca                                                  3200


<210>  7
<211>  633
<212>  DNA
<213>  Homo sapiens

<400>  7
actcaatgtc tgggctccaa accatgcata atatccacaa ctaaaatggc catgtcacca       60
agagagcttc ctctatttct cagattactg gaagattcat gcccaggagt atccataatt      120
tagcattcct gggatccgga cattctctct atcccaattt ttaatcatct gagtctgtat      180
taactgcttg ccagcagtta atacagactc agcattgtga cacttcccaa gaaggccaaa      240
cctctaaccg acattcctga aatacgtggc attattcttt tttggatttc tcatttatgg      300
aaggctaacc ctctgttgac cgtaagcctt ttggtttggg ctgtattgaa atcctttcta      360
aattgcatga ataggctctg ctaacgtgat gagacaaact gaaaattatt gcaagcattg      420
actataatta tgcagtacgt tctcaggatg catccagggg ttcattttca tgagcctgtc      480
caggttagtt tactcctgac cactaatagc attgtcattt gggctttctg ttgaatgaat      540
caacaaacca caatacttcc tgggaccttt tgtactttat ttgaactatg agtctgtatt      600
aactgcttgg ggatccttag agtcgagcca gca                                   633


<210>  8
<211>  377
<212>  DNA
<213>  Homo sapiens

<400>  8
tttgctgcta gtggttttga ggacctagcc ggctttattg cttatatttt gtatttagtt       60
tgtatttaac ctctagatct ttgactactg tgttgtcttt gactacggtt tcaaaatgta      120
tgtttttcaa tgtggagcac cccattcatg cattttttca ccctgtgcct tttacattta      180
taggtgtatt tgtcaaattt tgcatgagaa gatcacaatt ttagtgactc atcagttgca      240
gtacctcaaa gctgcaagtc agattctgat attgaaagat gtaagtagtt tctagaggtc      300
tcccgaactt gcaagtactt gggtgtgttg aatcccttc ttcttcctcc atttcatgtt       360
cctttcttct cagagct                                                     377


<210>  9
<211>  2602
<212>  DNA
<213>  Homo sapiens

<400>  9
cctccacctc ctgggttcaa gtgattctcc tgtctcagcc tcccgagtag ctggaattac       60
aggcgcaccc caccctgcct ggctactttt tgtattttta gtagagatgg gatttcacca      120
tgttggccag ctggtcttg aacttctgac caccagtgaa ccacctgcct cggcctccca      180
```

```
aagtgctggg attacaggca tgagccacca cacccagcct ggtcacatgt ttaatatcag    240
cttgtcttat gaacctgttt tgttcactgg aacaatgagg aaaaacctgg atccctttaa    300
tgagcacacg gatgaggaac tgtggaatgc cttacaagag gtacaactta aagaaaccat    360
tgaagatctt cctggtaaaa tggatactga attagcagaa tcaggatcca attttagtgt    420
tggacaaaga caactggtgt gccttgccag ggcaattctc aggaaaaatc agatattgat    480
tattgatgaa gcgacggcaa atgtggatcc aagaactgat gagttaatac aaaaaaaaat    540
ccgggagaaa tttgcccact gcaccgtgct aaccattgca cacagattga acaccattat    600
tgacagcgac aagataatgg ttttagattc aggaagactg aaagaatatg atgagccgta·    660
tgttttgctg caaaataaag agagcctatt ttacaagatg gtgcaacaac tgggcaaggc    720
agaagccgct gccctcactg aaacagcaaa acaggtatac ttcaaaagaa attatccaca    780
tattggtcac actgaccaca tggttacaaa cacttccaat ggacagccct cgaccttaac    840
tattttcgag acagcactgt gaatccaacc aaaatgtcaa gtccgttccg aaggcatttt    900
ccactagttt ttggactatg taaaccacat tgtacttttt tttactttgg caacaaatat    960
ttatacatac aagatgctag ttcatttgaa tatttctccc aacttatcca aggatctcca   1020
gctctaacaa aatggtttat ttttatttaa atgtcaatag ttgtttttta aaatccaaat   1080
cagaggtgca ggccaccagt taaatgccgt ctatcaggtt ttgtgcctta agagactaca   1140
gagtcaaagc tcattttttaa aggagtagga caaagttgtc acaggttttt gttgttgttt   1200
ttattgcccc caaaattaca tgttaatttc catttatatc agggattcta tttacttgaa   1260
gactgtgaag ttgccatttt gtctcattgt tttctttgac ataactagga tccattattt   1320
cccctgaagg cttcttgtta gaaaatagta cagttacaac caataggaac aacaaaaaga   1380
aaaagtttgt gacattgtag tagggagtgt gtaccccctta ctccccatca aaaaaaaaaa   1440
tggatacatg gttaaaggat agaagggcaa tattttatca tatgttctaa aagagaagga   1500
agagaaaata ctactttctc aaaatggaag cccttaaagg tgctttgata ctgaaggaca   1560
caaatgtgac cgtccatcct cctttagagt tgcatgactt ggacacggta actgttgcag   1620
ttttagactc agcattgtga cacttcccaa gaaggccaaa cctctaaccg acattcctga   1680
aatacgtggc attattcttt tttggatttc tcatttatgg aaggctaacc ctctgttgac   1740
cgtaagcctt ttggtttggg ctgtattgaa atcctttcta aattgcatga ataggctctg   1800
ctaacgtgat gagacaaact gaaaattatt gcaagcattg actataatta tgcagtacgt   1860
tctcaggatg catccagggg ttcattttca tgagcctgtc caggttagtt tactcctgac   1920
cactaatagc attgtcattt gggctttctg ttgaatgaat caacaaacca caatacttcc   1980
tgggaccttt tgtactttat ttgaactatg agtctttaat ttttcctgat gatggtggct   2040
gtaatatgtt gagttcagtt tactaaaggt tttactatta tggtttgaag tggagtctca   2100
tgacctctca gaataaggtg tcacctccct gaaattgcat atatgtatat agacatgcac   2160
acgtgtgcat ttgtttgtat acatatattt gtccttcgta tagcaagttt tttgctcatc   2220
agcagagagc aacagatgtt ttattgagtg aagccttaaa aagcacacac cacacacagc   2280
taactgccaa aatacattga ccgtagtagc tgttcaactc ctagtactta gaaatacacg   2340
tatggttaat gttcagtcca acaaaccaca cacagtaaat gtttattaat agtcatggtt   2400
cgtattttag gtgactgaaa ttgcaacagt gatcataatg aggtttgtta aaacgatagc   2460
tatattcaaa atgtctatat gtttatttgg acttttgagg ttaaagacag tcatataaac   2520
gtcctgtttc tgttttaatg ttatcataga attttttaat gaaactaaat tcaattgaaa   2580
taaatgatag ttttcatctc ca                                            2602
```

<210> 10
<211> 636
<212> DNA
<213> Homo sapiens

<400> 10
```
atggataagt ttatactagt gttggcacat ggcggcatgt atagatatay yaggassacc     60
takttgtatt ccttgtatgr aaaagcgtcc ctggtactac aataagtctt tcgwgaaagg    120
agtgtaatcc taacaacaac tcakgmaagt attttgaaaa gaatacttgg ataaggaaaa    180
acctgcakct actcctgcta tttcaagaca ttgctacaag tggttggtgt ggtctctgtg    240
gctgtggccg tgattccttg gatcgcaata cccttggttc cccttggaat cattttcatt    300
tttcttcggc gatatttttt ggaaacgtca agagatgtga agcgcctgga atctacaagt    360
gagtatggaa actcgggttg gtatagacat gctagctagt ttccatttat gccataaatt    420
acagagaccc cctgaaattc ggcagactct gtcttccaga atttctctaa cattaggtaa    480
ttgaacgtat tggccattat gaatcattgt gtcccttaga gcatgtggaa ttgatagcct    540
gcaacgtgta actttgcatt tggaataagg aaggagtgaa ggccatatgg ggagtaatat    600
```

                                9

tctacaggaa tgtcagcact gtgaagacag ggactc   636

<210>   11
<211>   2420
<212>   DNA
<213>   Homo sapiens

<400>   11
aagatcttat caatccaaat tttaccttcg ggttctgacg atttaaggaa gaaaatgtca   60
atcataccga acctgttttg ttcactggaa caatgaggaa aaacctggat ccctttaatg   120
agcacacgga tgaggaactg tggaatgcct acaagaggt acaacttaaa gaaaccattg   180
aagatcttcc tggtaaaatg gatactgaat tagcagaatc aggatccaat tttagtgttg   240
gacaaagaca actggtgtgc cttgccaggg caattctcag gaaaaatcag atattgatta   300
ttgatgaagc gacggcaaat gtggatccaa gaactgatga gttaatacaa aaaaaaatcc   360
gggagaaatt tgcccactgc accgtgctaa ccattgcaca cagattgaac accattattg   420
acagcgacaa gataatggtt ttagattcag gaagactgaa agaatatgat gagccgtatg   480
ttttgctgca aaataaagag agcctatttt acaagatggt gcaacaactg ggcaaggcag   540
aagccgctgc cctcactgaa acagcaaaac aggtatactt caaaagaaat tatccacata   600
ttggtcacac tgaccacatg gttacaaaca cttccaatgg acagccctcg accttaacta   660
tttttcgagac agcactgtga atccaaccaa aatgtcaagt ccgttccgaa ggcatttttcc   720
actagttttt ggactatgta aaccacattg tactttttt tactttggca acaaatattt   780
atacatacaa gatgctagtt catttgaata tttctcccaa cttatccaag gatctccagc   840
tctaacaaaa tggtttattt ttatttaaat gtcaatagtt gttttttaaa atccaaatca   900
gaggtgcagg ccaccagtta aatgccgtct atcaggtttt gtgccttaag agactacaga   960
gtcaaagctc atttttaaag gagtaggaca aagttgtcac aggttttgt tgttgttttt   1020
attgcccccca aaattacatg ttaatttcca tttatatcag ggattctatt tacttgaaga   1080
ctgtgaagtt gccattttgt ctcattgttt tctttgacat aactaggatc cattatttcc   1140
cctgaaggct tcttgttaga aaatagtaca gttacaacca ataggaacaa caaaaagaaa   1200
aagtttgtga cattgtagta gggagtgtgt accccttact ccccatcaaa aaaaaaaatg   1260
gatacatggt taaaggatag aagggcaata ttttatcata tgttctaaaa gagaaggaag   1320
agaaaatact actttctcaa aatggaagcc cttaaaggtg ctttgatact gaaggacaca   1380
aatgtgaccg tccatcctcc tttagagttg catgacttgg acacggtaac tgttgcagtt   1440
ttagactcag cattgtgaca cttcccaaga aggccaaacc tctaaccgac attcctgaaa   1500
tacgtggcat tattcttttt tggatttctc atttatggaa ggctaaccct ctgttgaccg   1560
taagccttt ggtttgggct gtattgaaat cctttctaaa ttgcatgaat aggctctgct   1620
aacgtgatga gacaaactga aaattattgc aagcattgac tataattatg cagtacgttc   1680
tcaggatgca tccaggggtt cattttcatg agcctgtcca ggttagttta ctcctgacca   1740
ctaatagcat tgtcatttgg gctttctgtt gaatgaatca acaaaccaca atacttcctg   1800
ggaccttttg tactttattt gaactatgag tctttaattt ttcctgatga tggtggctgt   1860
aatatgttga gttcagttta ctaaaggttt tactattatg gtttgaagtg gagtctcatg   1920
acctctcaga ataaggtgtc acctccctga aattgcatat atgtatatag acatgcacac   1980
gtgtgcattt gtttgtatac atatatttgt ccttcgtata gcaagttttt tgctcatcag   2040
cagagagcaa cagatgtttt attgagtgaa gccttaaaaa gcacacacca cacacagcta   2100
actgccaaaa tacattgacc gtagtagctg ttcaactcct agtacttaga aatacacgta   2160
tggttaatgt tcagtccaac aaaccacaca cagtaaatgt ttattaatag tcatggttcg   2220
tattttaggt gactgaaatt gcaacagtga tcataatgag gtttgttaaa acgatagcta   2280
tattcaaaat gtctatatgt ttatttggac ttttgaggtt aaagacagtc atataaacgt   2340
cctgtttctg ttttaatgtt atcatagaat tttttaatga aactaaattc aattgaaata   2400
aatgatagtt ttcatctcca   2420

<210>   12
<211>   703
<212>   DNA
<213>   Homo sapiens

<400>   12
caagcagggc atgagttcac caccccagta acaacagtgc attacaatac ttaataatgg   60

```
aacaaggaga gtagtatttt agtctcaaaa gcattggcac ttacaaccta ttgctttgtt      120
tttagtggtt ccagctgtg gttgcattgc acgtagaaag tggaataatg taatgggctt      180
tgaaaccata ataatgaatg tctgaataat gacattattt cttgcgtttg taatactgtt      240
aattaaatct atgtcgatcc tgttggaatt cataaaatca tctaaaaatt tttctaaata      300
tacagtgttg ttttccccat tgtatcttga tctcaagcaa caaatggtaa aagtatagct      360
attaatgtca ttaaatgtga attgtttcaa cattatgaag ggttcctctt ggtaagtggc      420
agaaggagcc aggcttaggt ttgaagtgag actgactttta ttcccttctt cctccttacg      480
tagctctgga tgccgggcag gttggtttgg cactgtccta tgccctcacg ctcatgggga      540
tgtttcagtg gtgtgttcga caaagtgctg aagttgagaa tatgatgcct gtaatcccag      600
ctactagaga ggctgaggca ggagaatcgc ttgaacctgg gaggcggagg ttgtggtgag      660
ccgagatctc acaattgcac tccagcctgg gcaacaagag tga                       703
```

```
<210>   13
<211>   1688
<212>   DNA
<213>   Homo sapiens

<400>   13
tcacagaaag ccacggtacc attgtgtatg tgtatcgtat attgcctgta tcacttagct       60
ttctggaatc taaaggttcc ttgggtgtgc aactttggtg ccctactcca cgtaaaaatc      120
ttctgtagct caactgagca aaacacactc gcacattcat ctacttattc atcatgctta      180
ttaccgttct caaagtctat caggttttgt gccttaagag actacagagt caaagctcat      240
ttttaaagga gtaggacaaa gttgtcacag gttttgttg ttgtttttat tgccccccaaa      300
attacatgtt aatttccatt tatatcaggg attctattta cttgaagact gtgaagttgc      360
cattttgtct cattgttttc tttgacataa ctaggatcca ttatttcccc tgaaggcttc      420
ttgttagaaa atagtacagt tacaaccaat aggaacaaca aaaagaaaaa gtttgtgaca      480
ttgtagtagg gagtgtgtac cccttactcc ccatcaaaaa aaaaaatgga tacatggtta      540
aaggatagaa gggcaatatt ttatcatatg ttctaaaaga gaaggaagag aaaatactac      600
tttctcaaaa tggaagccct taaaggtgct ttgatactga aggacacaaa tgtgaccgtc      660
catcctcctt tagagttgca tgacttggac acggtaactg ttgcagtttt agactcagca      720
ttgtgacact tcccaagaag gccaaacctc taaccgacat tcctgaaata cgtggcatta      780
ttctttttg gatttctcat ttatggaagg ctaaccctct gttgaccgta agcctttgg      840
tttgggctgt attgaaatcc tttctaaatt gcatgaatag gctctgctaa cgtgatgaga      900
caaactgaaa attattgcaa gcattgacta taattatgca gtacgttctc aggatgcatc      960
caggggttca ttttcatgag cctgtccagg ttagtttact cctgaccact aatagcattg     1020
tcatttgggc tttctgttga atgaatcaac aaaccacaat acttcctggg accttttgta     1080
ctttatttga actatgagtc tttaattttt cctgatgatg gtggctgtaa tatgttgagt     1140
tcagtttact aaaggtttta ctattatggt ttgaagtgga gtctcatgac ctctcagaat     1200
aaggtgtcac ctccctgaaa ttgcatatat gtatatagac atgcacagt gtgcatttgt     1260
ttgtatacat atatttgtcc ttcgtatagc aagttttttg ctcatcagca gagagcaaca     1320
gatgttttat tgagtgaagc cttaaaaagc acacaccaca cacagctaac tgccaaaata     1380
cattgaccgt agtagctgtt caactcctag tacttagaaa tacacgtatg gttaatgttc     1440
agtccaacaa accacacaca gtaaatgttt attaatagtc atggttcgta ttttaggtga     1500
ctgaaattgc aacagtgatc ataatgaggt ttgttaaaac gatagctata ttcaaaatgt     1560
ctatatgttt atttggactt ttgaggttaa agacagtcat ataaacgtcc tgtttctgtt     1620
ttaatgttat catagaattt tttaatgaaa ctaaattcaa ttgaaataaa tgatagtttt     1680
catctcca                                                             1688
```

```
<210>   14
<211>   5833
<212>   DNA
<213>   Homo sapiens

<400>   14.
ggacaggcgt ggcggccgga gccccagcat ccctgcttga ggtccaggag cggagcccgc       60
ggccaccgcc gcctgatcag cgcgaccccg gcccgcgccc gccccgcccg gcaagatgct      120
gcccgtgtac caggaggtga agcccaaccc gctgcaggac gcgaacatct gctcacgcgt      180
```

```
gttcttctgg tggctcaatc ccttgtttaa aattggccat aaacggagat tagaggaaga   240
tgatatgtat tcagtgctgc cagaagaccg ctcacagcac cttggagagg agttgcaagg   300
gttctgggat aaagaagttt taagagctga gaatgacgca cagaagcctt ctttaacaag   360
agcaatcata aagtgttact ggaaatctta tttagttttg ggaattttta cgttaattga   420
ggaaagtgcc aaagtaatcc agcccatatt tttgggaaaa attattaatt attttgaaaa   480
ttatgatccc atggattctg tggctttgaa cacagcgtac gcctatgcca cggtgctgac   540
tttttgcacg ctcattttgg ctatactgca tcacttatat ttttatcacg ttcagtgtgc   600
tgggatgagg ttacgagtag ccatgtgcca tatgatttat cggaaggcac ttcgtcttag   660
taacatggcc atggggaaga caaccacagg ccagatagtc aatctgctgt ccaatgatgt   720
gaacaagttt gatcaggtga cagtgttctt acacttcctg tgggcaggac cactgcaggc   780
gatcgcagtg actgccctac tctggatgga gataggaata tcgtgccttg ctgggatggc   840
agttctaatc attctcctgc ccttgcaaag ctgttttggg aagttgttct catcactgag   900
gagtaaaact gcaactttca cggatgccag gatcaggacc atgaatgaag ttataactgg   960
tataaggata ataaaaatgt acgcctggga aaagtcattt tcaaatctta ttaccaattt  1020
gagaaagaag gagatttcca agattctgag aagttcctgc ctcaggggga tgaatttggc  1080
ttcgttttc agtgcaagca aaatcatcgt gtttgtgacc ttcaccacct acgtgctcct  1140
cggcagtgtg atcacagcca gccgcgtgtt cgtggcagtg acgctgtatg gggctgtgcg  1200
gctgacggtt accctcttct tcccctcagc cattgagagg gtgtcagagg caatcgtcag  1260
catccgaaga atccagacct ttttgctact tgatgagata tcacagcgca accgtcagct  1320
gccgtcagat ggtaaaaaga tggtgcatgt gcaggatttt actgcttttt gggataaggc  1380
atcagagacc ccaactctac aaggcctttc ctttactgtc agacctggcg aattgttagc  1440
tgtggtcggc cccgtgggag cagggaagtc atcactgtta agtgccgtgc tcggggaatt  1500
ggccccaagt cacgggctgg tcagcgtgca tggaagaatt gcctatgtgt ctcagcagcc  1560
ctgggtgttc tcgggaactc tgaggagtaa tatttattt gggaagaaat atgaaaagga  1620
acgatatgaa aaagtcataa aggcttgtgc tctgaaaaag gatttacagc tgttggagga  1680
tggtgatctg actgtgatag gagatcgggg aaccacgctg agtggagggc agaaagcacg  1740
ggtaaacctt gcaagagcag tgtatcaaga tgctgacatc tatctcctgg acgatcctct  1800
cagtgcagta gatgcggaag ttagcagaca cttgttcgaa ctgtgtattt gtcaaatttt  1860
gcatgagaag atcacaattt tagtgactca tcagttgcag tacctcaaag ctgcaagtca  1920
gattctgata ttgaaagatg gtaaaatggt gcagaagggg acttacactg agttcctaaa  1980
atctggtata gattttggct cccttttaaa gaaggataat gaggaaagtg aacaacctcc  2040
agttccagga actcccacac taaggaatcg taccttctca gagtcttcgg tttggtctca  2100
acaatcttct agaccctcct tgaaagatgg tgctctggag agccaagata cagagaatgt  2160
cccagttaca ctatcagagg agaaccgttc tgaaggaaaa gttggttttc aggcctataa  2220
gaattacttc agagctggtg ctcactggat tgtcttcatt ttccttattc tcctaaacac  2280
tgcagctcag gttgcctatg tgcttcaaga ttggtggctt tcatactggg caaacaaaca  2340
aagtatgcta aatgtcactg taaatggagg aggaaatgta accgagaagc tagatcttaa  2400
ctggtactta ggaatttatt caggtttaac tgtagctacc gttcttttg gcatagcaag  2460
atctctattg gtattctacg tccttgttaa ctcttcacaa actttgcaca caaaatgtt  2520
tgagtcaatt ctgaaagctc cggtattatt ctttgataga aatccaatag gaagaatttt  2580
aaatcgtttc tccaaagaca ttggacactt ggatgatttg ctgccgctga cgtttttaga  2640
tttcatccag acattgctac aagtggttgg tgtggtctct gtggctgtgg ccgtgattcc  2700
ttggatcgca atacccttgg ttccccttgg aatcatttc atttttcttc ggcgatattt  2760
tttggaaacg tcaagagatg tgaagcgcct ggaatctaca actcggagtc cagtgttttc  2820
ccacttgtca tcttctctcc aggggctctg gaccatccgg catacaaag cagaagagag  2880
gtgtcaggaa ctgtttgatg cacaccagga tttacattca gaggcttggt tcttgttttt  2940
gacaacgtcc cgctggttcg ccgtccgtct ggatgccatc tgtgccatgt tgtcatcat  3000
cgttgccttt gggtccctga ttctggcaaa aactctggat gccgggcagg ttggtttggc  3060
actgtcctat gccctcacgc tcatggggat gtttcagtgg tgtgttcgac aaagtgctga  3120
agttgagaat atgatgatct cagtagaaag ggtcattgaa tacacagacc ttgaaaaaga  3180
agcaccttgg gaatatcaga acgcccacc accagcctgg ccccatgaag gagtgataat  3240
ctttgacaat gtgaacttca gtgtacagtcc aggtgggcct ctggtactga agcatctgac  3300
agcactcatt aaatcacaag aaaggttgg cattgtggga agaaccggag ctggaaaaag  3360
ttccctcatc tcagcccttt ttagattgtc agaacccgaa ggtaaaattt ggattgataa  3420
gatcttgaca actgaaattg gacttcacga tttaaggaag aaaatgtcaa tcatacctca  3480
ggaacctgtt ttgttcactg gaacaatgag gaaaaacctg gatcccttta atgagcacac  3540
ggatgaggaa ctgtggaatg ccttacaaga ggtacaactt aaagaaacca ttgaagatct  3600
tcctggtaaa atggatactg aattagcaga atcaggatcc aatttttagtg ttggacaaag  3660
acaactggtg tgccttgcca gggcaattct caggaaaaat cagatattga ttattgatga  3720
```

12

```
agcgacggca aatgtggatc caagaactga tgagttaata caaaaaaaaa tccgggagaa   3780
atttgcccac tgcaccgtgc taaccattgc acacagattg aacaccatta ttgacagcga   3840
caagataatg gttttagatt caggaagact gaaagaatat gatgagccgt atgttttgct   3900
gcaaaataaa gagagcctat tttacaagat ggtgcaacaa ctgggcaagg cagaagccgc   3960
tgccctcact gaaacagcaa aacaggtata cttcaaaaga aattatccac atattggtca   4020
cactgaccac atggttacaa acacttccaa tggacagccc tcgaccttaa ctattttcga   4080
gacagcactg tgaatccaac caaaatgtca agtccgttcc gaaggcattt tccactagtt   4140
tttggactat gtaaaccaca ttgtactttt ttttactttg gcaacaaata tttatacata·  4200
caagatgcta gttcatttga atatttctcc caacttatcc aaggatctcc agctctaaca   4260
aaatggttta tttttattta aatgtcaata gttgtttttt aaaatccaaa tcagaggtgc   4320
aggccaccag ttaaatgccg tctatcaggt tttgtgcctt aagagactac agagtcaaag   4380
ctcattttta aaggagtagg acaaagttgt cacaggtttt tgttgttgtt tttattgccc   4440
ccaaaattac atgttaattt ccatttatat cagggattct atttacttga agactgtgaa   4500
gttgccattt tgtctcattg ttttctttga cataactagg atccattatt tcccctgaag   4560
gcttcttgtt agaaaatagt acagttacaa ccaataggaa caacaaaaag aaaaagtttg   4620
tgacattgta gtagggagtg tgtacccctt actccccatc aaaaaaaaaa atggatacat   4680
ggttaaagga tagaagggca atattttatc atatgttcta aaagagaagg aagagaaaat   4740
actactttct caaaatggaa gcccttaaag gtgctttgat actgaaggac acaaatgtga   4800
ccgtccatcc tcctttagag ttgcatgact tggacacggt aactgttgca gttttagact   4860
cagcattgtg acacttccca agaaggccaa acctctaacc gacattcctg aaatacgtgg   4920
cattattctt ttttggattt ctcatttatg gaaggctaac cctctgttga ccgtaagcct   4980
tttggtttgg gctgtattga aatcctttct aaattgcatg aataggctct gctaacgtga   5040
tgagacaaac tgaaaattat tgcaagcatt gactataatt atgcagtacg ttctcaggat   5100
gcatccaggg gttcattttc atgagcctgt ccaggttagt ttactcctga ccactaatag   5160
cattgtcatt tgggctttct gttgaatgaa tcaacaaacc acaatacttc ctgggacctt   5220
ttgtacttta tttgaactat gagtctttaa ttttcctga tgatggtggc tgtaatatgt   5280
tgagttcagt ttactaaagg ttttactatt atggtttgaa gtggagtctc atgacctctc   5340
agaataaggt gtcacctccc tgaaattgca tatatgtata tagacatgca cacgtgtgca   5400
tttgtttgta tacatatatt tgtccttcgt atagcaagtt ttttgctcat cagcagagag   5460
caacagatgt tttattgagt gaagccttaa aaagcacaca ccacacacag ctaactgcca   5520
aaatacattg accgtagtag ctgttcaact cctagtactt agaaatacac gtatggttaa   5580
tgttcagtcc aacaaaccac acacagtaaa tgtttattaa tagtcatggt tcgtatttta   5640
ggtgactgaa attgcaacag tgatcataat gaggtttgtt aaaacgatag ctatattcaa   5700
aatgtctata tgtttatttg gacttttgag gttaaagaca gtcatataaa cgtcctgttt   5760
ctgtttaat gttatcatag aattttttaa tgaaactaaa ttcaattgaa ataaatgata   5820
gttttcatct cca                                                     5833
```

```
<210>   15
<211>   1182
<212>   DNA
<213>   Homo sapiens

<400>   15
actcaatgtc tgggctccaa accatgcata atatccacaa ctaaaatggc catgtcacca    60
agagagcttc ctctatttct cagattactg gaagattcat gcccaggagt atccataatt   120
tagcattcct gggatccgga cattctctct atcccaattt ttaatcatct gagtctgtat   180
taactgcttg ccagcagtta atacagactc agcattgtga cacttcccaa gaaggccaaa   240
cctctaaccg acattcctga aatacgtggc attattcttt tttggatttc tcatttatgg   300
aaggctaacc ctctgttgac cgtaagcctt ttggtttggg ctgtattgaa atcctttcta   360
aattgcatga ataggctctg ctaacgtgat gagacaaact gaaaattatt gcaagcattg   420
actataatta tgcagtacgt tctcaggatg catccagggg ttcattttca tgagcctgtc   480
caggttagtt tactcctgac cactaatagc attgtcattt gggctttctg ttgaatgaat   540
caacaaacca caatacttcc tgggaccttt gtactttat ttgaactatg agtctttaat   600
ttttcctgat gatggtggct gtaatatgtt gagttcagtt tactaaaggt tttactatta   660
tggtttgaag tggagtctca tgacctctca gaataaggtg tcacctccct gaaattgcat   720
atatgtatat agacatgcac acgtgtgcat ttgtttgtat acatatattt gtccttcgta   780
tagcaagttt tttgctcatc agcagagagc aacagatgtt ttattgagtg aagccttaaa   840
aagcacacac cacacacagc taactgccaa aatacattga ccgtagtagc tgttcaactc   900
```

```
ctagtactta  gaaatacacg  tatggttaat  gttcagtcca  acaaaccaca  cacagtaaat        960
gtttattaat  agtcatggtt  cgtattttag  gtgactgaaa  ttgcaacagt  gatcataatg       1020
aggtttgtta  aaacgatagc  tatattcaaa  atgtctatat  gtttatttgg  acttttgagg       1080
ttaaagacag  tcatataaac  gtcctgtttc  tgttttaatg  ttatcataga  attttttaat       1140
gaaactaaat  tcaattgaaa  taaatgatag  ttttcatctc  ca                          1182
```

```
<210>   16
<211>   5862
<212>   DNA
<213>   Homo sapiens

<400>   16
ggacaggcgt' ggcggccgga  gccccagcat  ccctgcttga  ggtccaggag  cggagcccgc         60
ggccaccgcc  gcctgatcag  cgcgaccccg  gcccgcgccc  gccccgcccg  gcaagatgct        120
gcccgtgtac  caggaggtga  agcccaaccc  gctgcaggac  gcgaacatct  gctcacgcgt        180
gttcttctgg  tggctcaatc  ccttgtttaa  aattggccat  aaacggagat  tagaggaaga        240
tgatatgtat  tcagtgctgc  cagaagaccg  ctcacagcac  cttggagagg  agttgcaagg        300
gttctgggat  aaagaagttt  taagagctga  gaatgacgca  cagaagcctt  ctttaacaag        360
agcaatcata  aagtgttact  ggaaatctta  tttagttttg  ggaattttta  cgttaattga        420
ggaaagtgcc  aaagtaatcc  agcccatatt  tttgggaaaa  attattaatt  attttgaaaa        480
ttatgatccc  atggattctg  tggctttgaa  cacagcgtac  gcctatgcca  cggtgctgac        540
ttttgcacg  ctcattttgg  ctatactgca  tcacttatat  ttttatcacg  ttcagtgtgc        600
tgggatgagg  ttacgagtag  ccatgtgcca  tatgatttat  cggaaggcac  ttcgtcttag        660
taacatggcc  atggggaaga  caaccacagg  ccagatagtc  aatctgctgt  ccaatgatgt        720
gaacaagttt  gatcaggtga  cagtgttctt  acacttcctg  tgggcaggac  cactgcaggc        780
gatcgcagtg  actgccctac  tctggatgga  gataggaata  tcgtgccttg  ctgggatggc        840
agttctaatc  attctcctgc  ccttgcaaag  ctgttttggg  aagttgttct  catcactgag        900
gagtaaaact  gcaactttca  cggatgccag  gatcaggacc  atgaatgaag  ttataactgg        960
tataaggata  ataaaaatgt  acgcctggga  aaagtcattt  tcaaatctta  ttaccaattt       1020
gagaaagaag  gagatttcca  agattctgag  aagttcctgc  ctcaggggga  tgaatttggc       1080
ttcgttttc  agtgcaagca  aaatcatcgt  gtttgtgacc  ttcaccacct  acgtgctcct       1140
cggcagtgtg  atcacagcca  gccgcgtgtt  cgtggcagtg  acgctgtatg  gggctgtgcg       1200
gctgacggtt  accctcttct  tcccctcagc  cattgagagg  gtgtcagagg  caatcgtcag       1260
catccgaaga  atccagacct  ttttgctact  tgatgagata  tcacagcgca  accgtcagct       1320
gccgtcagat  ggtaaaaaga  tggtgcatgt  gcaggatttt  actgcttttt  gggataaggc       1380
atcagagacc  ccaactctac  aaggcctttc  ctttactgtc  agacctggcg  aattgttagc       1440
tgtggtcggc  cccgtgggag  cagggaagtc  atcactgtta  agtgccgtgc  tcggggaatt       1500
ggccccaagt  cacgggctgg  tcagcgtgca  tggaagaatt  gcctatgtgt  ctcagcagcc       1560
ctgggtgttc  tcgggaactc  tgaggagtaa  tattttattt  gggaagaaat  atgaaaagga       1620
acgatatgaa  aaagtcataa  aggcttgtgc  tctgaaaaag  gatttacagc  tgttggagga       1680
tggtgatctg  actgtgatag  gagatcgggg  aaccacgctg  agtggagggc  agaaagcacg       1740
ggtaaacctt  gcaagagcag  tgtatcaaga  tgctgacatc  tatctcctgg  acgatcctct       1800
cagtgcagta  gatgcggaag  ttagcagaca  cttgttcgaa  ctgtgtattt  gtcaaatttt       1860
gcatgagaag  atcacaattt  tagtgactca  tcagttgcag  tacctcaaag  ctgcaagtca       1920
gattctgata  ttgaaagatg  gtaaaatggt  gcagaagggg  acttacactg  agttcctaaa       1980
atctggtata  gatttтggct  cccttttaaa  gaaggataat  gaggaaagtg  aacaacctcc       2040
agttccagga  actcccacac  taaggaatcg  taccttctca  gagtcttcgg  tttggtctca       2100
acaatcttct  agaccctcct  tgaaagatgg  tgctctggag  agccaagata  cagagaatgt       2160
cccagttaca  ctatcagagg  agaaccgttc  tgaaggaaaa  gttggttttc  aggcctataa       2220
gaattacttc  agagctggtg  ctcactggat  tgtcttcatt  ttccttattc  tcctaaacac       2280
tgcagctcag  gttgcctatg  tgcttcaaga  ttggtggctt  tcatactggg  caaacaaaca       2340
aagtatgcta  aatgtcactg  taaatggagg  aggaaatgta  accgagaagc  tagatcttaa       2400
ctggtactta  ggaatttatt  caggtttaac  tgtagctacc  gttctttttg  gcatagcaag       2460
atctctattg  gtattctacg  tccttgttaa  ctcttcacaa  actttgcaca  acaaaatgtt       2520
tgagtcaatt  ctgaaagctc  cggtattatt  ctttgataga  aatccaatag  gaagaatttt       2580
aaatcgtttc  tccaaagaca  ttggacactt  ggatgatttg  ctgccgctga  cgttttttaga       2640
tttcatccag  acattgctac  aagtggttgg  tgtggtctct  gtggctgtgg  ccgtgattcc       2700
ttggatcgca  ataccccttgg  ttccccttgg  aatcattttc  atttttcttc  ggcgatattt       2760
```

```
tttggaaacg tcaagagatg tgaagcgcct ggaatctaca actcggagtc cagtgttttc   2820
ccacttgtca tcttctctcc aggggctctg gaccatccgg gcatacaaag cagaagagag   2880
gtgtcaggaa ctgtttgatg cacaccagga tttacattca gaggcttggt tcttgttttt   2940
gacaacgtcc cgctggttcg ccgtccgtct ggatgccatc tgtgccatgt ttgtcatcat   3000
cgttgccttt gggtccctga ttctggcaaa aactctggat gccgggcagg ttggtttggc   3060
actgtcctat gccctcacgc tcatggggat gtttcagtgg tgtgttcgac aaagtgctga   3120
agttgagaat atgatgatct cagtagaaag ggtcattgaa tacacagacc ttgaaaaaga   3180
agcaccttgg gaatatcaga aacgcccacc accagcctgg ccccatgaag gagtgataat   3240
ctttgacaat gtgaacttca tgtacagtcc aggtgggcct ctggtactga agcatctgac   3300
agcactcatt aaatcacaag aaaaggttgg cattgtggga agaaccggag ctggaaaaag   3360
ttccctcatc tcagcccttt ttagattgtc agaacccgaa ggtaaaattt ggattgataa   3420
gatcttgaca actgaaattg gacttcacga tttaaggaag aaaatgtcaa tcatacctca   3480
ggaacctgtt ttgttcactg gaacaatgag gaaaaacctg gatcccttta atgagcacac   3540
ggatgaggaa ctgtggaatg ccttacaaga ggtacaactt aaagaaacca ttgaagatct   3600
tcctggtaaa atggatactg aattagcaga atcaggatcc aatttttagtg ttggacaaag   3660
acaactggtg tgccttgcca gggcaattct caggaaaaat cagatattga ttattgatga   3720
agcgacggca aatgtggatc caagaactga tgagttaata caaaaaaaaa tccgggagaa   3780
atttgcccac tgcaccgtgc taaccattgc acacagattg aacaccatta ttgacagcga   3840
caagataatg gttttagatt caggaagact gaaagaatat gatgagccgt atgtttttgct  3900
gcaaaataaa gagagcctat tttacaagat ggtgcaacaa ctgggcaagg cagaagccgc   3960
tgccctcact gaaacagcaa aacaggtata cttcaaaaga aattatccac atattggtca   4020
cactgaccac atggttacaa acacttccaa tggacagccc tcgaccttaa ctattttcga   4080
gacagcactg tgaatccaac caaaatgtca agtccgttcc gaaggcattt tccactagtt   4140
tttggactat gtaaaccaca ttgtactttt ttttactttg gcaacaaata tttatacata   4200
caagatgcta gttcatttga atatttctcc caacttatcc aaggatctcc agctctaaca   4260
aaatggttta tttttattta aatgtcaata gttgtttttt aaaatccaaa tcagaggtgc   4320
aggccaccag ttaaatgccg tctatcaggt tttgtgcctt aagagactac agagtcaaag   4380
ctcatttta aaggagtagg acaaagttgt cacaggtttt tgttgttgtt tttattgccc   4440
ccaaaattac atgttaattt ccatttatat cagggattct atttacttga agactgtgaa   4500
gttgccattt tgtctcattg ttttctttga cataactagg atccattatt tcccctgaag   4560
gcttcttgtt agaaaatagt acagttacaa ccaataggaa caacaaaaag aaaaagtttg   4620
tgacattgta gtagggagtg tgtaccoctt actccccatc aaaaaaaaaa atggatacat   4680
ggttaaagga tagaagggca atatttatc atatgttcta aaagagaagg aagagaaat    4740
actactttct caaaatggaa gcccttaaag gtgctttgat actgaaggac acaaatgtga   4800
ccgtccatcc tcctttagag ttgcatgact tggacacggt aactgttgca gttttagact   4860
cagcattgtg acacttccca agaaggccaa acctctaacc gacattcctg aaatacgtgg   4920
cattattctt ttttggattt ctcatttatg gaaggctaac cctctgttga ccgtaagcct   4980
tttggtttgg gctgtattga aatcctttct aaattgcatg aataggctct gctaacgtga   5040
tgagacaaac tgaaaattat tgcaagcatt gactataatt atgcagtacg ttctcaggat   5100
gcatccaggg gttcattttc atgagcctgt ccaggttagt ttactcctga ccactaatag   5160
cattgtcatt tgggctttct gttgaatgaa tcaacaaacc acaatacttc ctgggacctt   5220
ttgtacttta tttgaactat gagtctttaa ttttttcctga tgatggtggc tgtaatatgt   5280
tgagttcagt ttactaaagg ttttactatt atggtttgaa gtggagtctc atgacctctc   5340
agaataaggt gtcacctccc tgaaattgca tatatgtata tagacatgca cacgtgtgca   5400
tttgtttgta tacatatatt tgtccttcgt atagcaagtt ttttgctcat cagcagagag   5460
caacagatgt tttattgagt gaagccttaa aaagcacaca ccacacacag ctaactgcca   5520
aaatacattg accgtagtag ctgttcaact cctagtactt agaaatacac gtatggttaa   5580
tgttcagtcc aacaaaccac acacagtaaa tgtttattaa tagtcatggt tcgtatttta   5640
ggtgactgaa attgcaacag tgatcataat gaggtttgtt aaaacgatag ctatattcaa   5700
aatgtctata tgtttatttg gacttttgag gttacatgca gtgacattca ggtcgcgttt   5760
cggttgtcat gagatcagag ggttttttgat ccagacccac ccccgcccca ccccgcgcgc   5820
ctttgctgcc tcttccagag acccagcttg cagagcggcc gc                      5862
```

```
<210>  17
<211>  1118
<212>  DNA
<213>  Homo sapiens
```

```
<400>  17
taaactagtc caggttccca cagctcgtaa gctgctgtct cttgtctccc aggtctcagc      60
tcagctgtgt ctttgggaca tctttccaga ccattcactc tgtaggagcc acctaccgct     120
gagtctgtag cagttaatac agactcagca ttgtgacact tcccaagaag gccaaacctc     180
taaccgacat tcctgaaata cgtggcatta ttcttttttg gatttctcat ttatggaagg     240
ctaaccctct gttgaccgta agccttttgg tttgggctgt attgaaatcc tttctaaatt     300
gcatgaatag gctctgctaa cgtgatgaga caaactgaaa attattgcaa gcattgacta     360
taattatgca gtacgttctc aggatgcatc caggggttca tttttcatgag cctgtccagg     420
ttagtttact cctgaccact aatagcattg tcatttgggc tttctgttga atgaatcaac     480
aaaccacaat acttcctggg accttttgta ctttatttga actatgagtc tttaattttt     540
cctgatgatg gtggctgtaa tatgttgagt tcagtttact aaaggtttta ctattatggt     600
ttgaagtgga gtctcatgac ctctcagaat aaggtgtcac ctccctgaaa ttgcatatat     660
gtatatagac atgcacacgt gtgcatttgt ttgtatacat atatttgtcc ttcgtatagc     720
aagttttttg ctcatcagca gagagcaaca gatgttttat tgagtgaagc cttaaaaagc     780
acacaccaca cacagctaac tgccaaaata cattgaccgt agtagctgtt caactcctag     840
tacttagaaa tacacgtatg gttaatgttc agtccaacaa accacacaca gtaaatgttt     900
attaatagtc atggttcgta ttttaggtga ctgaaattgc aacagtgatc ataatgaggt     960
ttgttaaaac gatagctata ttcaaaatgt ctatatgttt atttggactt ttgaggttaa    1020
agacagtcat ataaa  cc tgtttctgtt ttaatgttat catagaattt tttaatgaaa    1080
ctaaattcaa ttgaaataaa tgatagtttt catctcca                           1118


<210>  18
<211>  1531
<212>  DNA
<213>  Homo sapiens

<220>
<221>  misc_feature
<222>  (1)..(1531)
<223>  n=a,t,g or c

<400>  18
caagcagggc atgagttcac caccccagta acaacagtgc attacaatac ttaataatgg      60
aacaaggaga gtagtatttt agtctcaaaa gcattggcac ttacaaccta ttgctttgtt     120
tttagtggtt ccagctgtg gttgcattgc acgtagaaag tggaataatg taatgggctt     180
tgaaaccata ataatgaatg tctgaataat gacattattt cttgcgtttg taatactgtt     240
aattaaatct atgtcgatcc tgttggaatt cataaaatca tctaaaaatt tttctaaata     300
tacagtgttg ttttccccat tgtatcttga tctcaagcaa caaatggtaa aagtatagct     360
attaatgtca ttaaatgtga attgtttcaa cattatgaag ggttcctctt ggtaagtggc     420
agaaggagcc aggcttaggt ttgaagtgag actgacttta ttcccttctt cctccttacg     480
tagctctgga tgccgggcag gttggtttgg cactgtccta tgccctcacg ctcatgggga     540
tgtttcagtg gtgtgttcga caaagtgctg aagttgagaa tatgatgatc tcagtagaaa     600
gggtcattga atacacagac cttgaaaaag aagcaccttg ggaatatcag aaacgcccac     660
caccagcctg gccccatgaa ggagtgataa tctttgacaa tgtgaacttc atgtacagtc     720
caggtgggcc tctggtactg aagcatctga cagcactcat taaatcacaa gaaaaggttg     780
gcattgtggg aagaaccgga gctggaaaaa gttccctcat ctcagccctt tttagattgt     840
cagaacccga aggtaaaatt tggattgata agatcttgac aactgaaatt ggacttcacg     900
atttaaggaa gaaaatgtca atcatacctc aggaacctgt tttgttcact ggaacaatga     960
ggaaaaacct ggatcccttt aatgagcaca cggatgagga actgtggaat gccttacaag    1020
aggtacaact taaagaaacc attgaagatc ttcctggtaa aatggatact gaattagcag    1080
aatcaggatc caattttagt gttggacaaa gacaactggt gtgccttgcc agggcaattc    1140
tcaggaaaaa tcagatattg attattgatg aagcgacggc aaatgtggat ccaagaactg    1200
atgagttaat acaaaaaaaa atccgggaga aatttgccca ctgcaccgtg ctaaccattg    1260
cacacagatt gaacaccatt attgacagcg acaagataat ggttttagat tcaggaagac    1320
tgaaagaata tgatgagccg tatgttttgc tgcaaaataa agagagccta ttttttacaa    1380
ggtgggtggc caaccacttg gggccagggc nggaggccgc ttgcccttca cttggaaacc    1440
ngccaaaccg ggtttttactt tcccaaaggg attttttnccc cctttttggg ttcaccantg    1500
gacccctggg ggttnccaac cctttcccctt t                                  1531
```

```
<210>   19
<211>   286
<212>   PRT
<213>   Homo sapiens

<400>   19
Leu His Leu Leu Gly Ser Ser Asp Ser Pro Val Ser Ala Ser Arg Val
1               5                   10                  15

Ala Gly Ile Thr Gly Ala Pro His Pro Ala Trp Leu Leu Phe Val Phe
            20                  25                  30

Leu Val Glu Met Gly Phe His His Val Gly Gln Ala Gly Leu Glu Leu
        35                  40                  45

Leu Thr Thr Ser Glu Pro Pro Ala Ser Ala Ser Gln Ser Ala Gly Ile
    50                  55                  60

Thr Gly Met Ser His His Thr Gln Pro Gly His Met Phe Asn Ile Ser
65                  70                  75                  80

Leu Ser Tyr Glu Pro Val Leu Phe Thr Gly Thr Met Arg Lys Asn Leu
            85                  90                  95

Asp Pro Phe Asn Glu His Thr Asp Glu Glu Leu Trp Asn Ala Leu Gln
            100                 105                 110

Glu Val Gln Leu Lys Glu Thr Ile Glu Asp Leu Pro Gly Lys Met Asp
        115                 120                 125

Thr Glu Leu Ala Glu Ser Gly Ser Asn Phe Ser Val Gly Gln Arg Gln
        130                 135                 140

Leu Val Cys Leu Ala Arg Ala Ile Leu Arg Lys Asn Gln Ile Leu Ile
145                 150                 155                 160

Ile Asp Glu Ala Thr Ala Asn Val Asp Pro Arg Thr Asp Glu Leu Ile
                165                 170                 175

Gln Lys Lys Ile Arg Glu Lys Phe Ala His Cys Thr Val Leu Thr Ile
            180                 185                 190

Ala His Arg Leu Asn Thr Ile Ile Asp Ser Asp Lys Ile Met Val Leu
        195                 200                 205

Asp Ser Gly Arg Leu Lys Glu Tyr Asp Glu Pro Tyr Val Leu Leu Gln
    210                 215                 220

Asn Lys Glu Ser Leu Phe Tyr Lys Met Val Gln Gln Leu Gly Lys Ala
225                 230                 235                 240

Glu Ala Ala Ala Leu Thr Glu Thr Ala Lys Gln Val Tyr Phe Lys Arg
                245                 250                 255

Asn Tyr Pro His Ile Gly His Thr Asp His Met Val Thr Asn Thr Ser
            260                 265                 270

Asn Gly Gln Pro Ser Thr Leu Thr Ile Phe Glu Thr Ala Leu
            275                 280                 285
```

```
<210>  20
<211>  225
<212>  PRT
<213>  Homo sapiens

<400>  20
Asp Leu Ile Asn Pro Asn Phe Thr Phe Gly Phe Trp Arg Phe Lys Glu
1               5                   10                  15

Glu Asn Val Asn His Thr Glu Pro Val Leu Phe Thr Gly Thr Met Arg
            20                  25                  30

Lys Asn Leu Asp Pro Phe Asn Glu His Thr Asp Glu Glu Leu Trp Asn
        35                  40                  45

Ala Leu Gln Glu Val Gln Leu Lys Glu Thr Ile Glu Asp Leu Pro Gly
    50                  55                  60

Lys Met Asp Thr Glu Leu Ala Glu Ser Gly Ser Asn Phe Ser Val Gly
65                  70                  75                  80

Gln Arg Gln Leu Val Cys Leu Ala Arg Ala Ile Leu Arg Lys Asn Gln
                85                  90                  95

Ile Leu Ile Ile Asp Glu Ala Thr Ala Asn Val Asp Pro Arg Thr Asp
            100                 105                 110

Glu Leu Ile Gln Lys Lys Ile Arg Glu Lys Phe Ala His Cys Thr Val
        115                 120                 125

Leu Thr Ile Ala His Arg Leu Asn Thr Ile Ile Asp Ser Asp Lys Ile
    130                 135                 140

Met Val Leu Asp Ser Gly Arg Leu Lys Glu Tyr Asp Glu Pro Tyr Val
145                 150                 155                 160

Leu Leu Gln Asn Lys Glu Ser Leu Phe Tyr Lys Met Val Gln Gln Leu
                165                 170                 175

Gly Lys Ala Glu Ala Ala Ala Leu Thr Glu Thr Ala Lys Gln Val Tyr
            180                 185                 190

Phe Lys Arg Asn Tyr Pro His Ile Gly His Thr Asp His Met Val Thr
        195                 200                 205

Asn Thr Ser Asn Gly Gln Pro Ser Thr Leu Thr Ile Phe Glu Thr Ala
    210                 215                 220

Leu
225


<210>  21
<211>  1325
<212>  PRT
<213>  Homo sapiens

<400>  21
```

Met Leu Pro Val Tyr Gln Glu Val Lys Pro Asn Pro Leu Gln Asp Ala
1   5   10   15

Asn Ile Cys Ser Arg Val Phe Phe Trp Trp Leu Asn Pro Leu Phe Lys
             20   25   30

Ile Gly His Lys Arg Arg Leu Glu Glu Asp Asp Met Tyr Ser Val Leu
             35   40   45

Pro Glu Asp Arg Ser Gln His Leu Gly Glu Glu Leu Gln Gly Phe Trp
    50   55   60

Asp Lys Glu Val Leu Arg Ala Glu Asn Asp Ala Gln Lys Pro Ser Leu
65   70   75   80

Thr Arg Ala Ile Ile Lys Cys Tyr Trp Lys Ser Tyr Leu Val Leu Gly
             85   90   95

Ile Phe Thr Leu Ile Glu Glu Ser Ala Lys Val Ile Gln Pro Ile Phe
         100   105   110

Leu Gly Lys Ile Ile Asn Tyr Phe Glu Asn Tyr Asp Pro Met Asp Ser
         115   120   125

Val Ala Leu Asn Thr Ala Tyr Ala Tyr Ala Thr Val Leu Thr Phe Cys
    130   135   140

Thr Leu Ile Leu Ala Ile Leu His His Leu Tyr Phe Tyr His Val Gln
145   150   155   160

Cys Ala Gly Met Arg Leu Arg Val Ala Met Cys His Met Ile Tyr Arg
             165   170   175

Lys Ala Leu Arg Leu Ser Asn Met Ala Met Gly Lys Thr Thr Thr Gly
         180   185   190

Gln Ile Val Asn Leu Leu Ser Asn Asp Val Asn Lys Phe Asp Gln Val
         195   200   205

Thr Val Phe Leu His Phe Leu Trp Ala Gly Pro Leu Gln Ala Ile Ala
    210   215   220

Val Thr Ala Leu Leu Trp Met Glu Ile Gly Ile Ser Cys Leu Ala Gly
225   230   235   240

Met Ala Val Leu Ile Ile Leu Leu Pro Leu Gln Ser Cys Phe Gly Lys
             245   250   255

Leu Phe Ser Ser Leu Arg Ser Lys Thr Ala Thr Phe Thr Asp Ala Arg
             260   265   270

Ile Arg Thr Met Asn Glu Val Ile Thr Gly Ile Arg Ile Ile Lys Met
             275   280   285

Tyr Ala Trp Glu Lys Ser Phe Ser Asn Leu Ile Thr Asn Leu Arg Lys
    290   295   300

Lys Glu Ile Ser Lys Ile Leu Arg Ser Ser Cys Leu Arg Gly Met Asn
305   310   315   320

Leu Ala Ser Phe Phe Ser Ala Ser Lys Ile Ile Val Phe Val Thr Phe
                325   330   335

Thr Thr Tyr Val Leu Leu Gly Ser Val Ile Thr Ala Ser Arg Val Phe
            340   345   350

Val Ala Val Thr Leu Tyr Gly Ala Val Arg Leu Thr Val Thr Leu Phe
            355   360   365

Phe Pro Ser Ala Ile Glu Arg Val Ser Glu Ala Ile Val Ser Ile Arg
    370   375   380

Arg Ile Gln Thr Phe Leu Leu Leu Asp Glu Ile Ser Gln Arg Asn Arg
385   390   395   400

Gln Leu Pro Ser Asp Gly Lys Lys Met Val His Val Gln Asp Phe Thr
            405   410   415

Ala Phe Trp Asp Lys Ala Ser Glu Thr Pro Thr Leu Gln Gly Leu Ser
            420   425   430

Phe Thr Val Arg Pro Gly Glu Leu Leu Ala Val Val Gly Pro Val Gly
    435   440   445

Ala Gly Lys Ser Ser Leu Leu Ser Ala Val Leu Gly Glu Leu Ala Pro
    450   455   460

Ser His Gly Leu Val Ser Val His Gly Arg Ile Ala Tyr Val Ser Gln
465   470   475   480

Gln Pro Trp Val Phe Ser Gly Thr Leu Arg Ser Asn Ile Leu Phe Gly
            485   490   495

Lys Lys Tyr Glu Lys Glu Arg Tyr Glu Lys Val Ile Lys Ala Cys Ala
            500   505   510

Leu Lys Lys Asp Leu Gln Leu Leu Glu Asp Gly Asp Leu Thr Val Ile
    515   520   525

Gly Asp Arg Gly Thr Thr Leu Ser Gly Gly Gln Lys Ala Arg Val Asn
    530   535   540

Leu Ala Arg Ala Val Tyr Gln Asp Ala Asp Ile Tyr Leu Leu Asp Asp
545   550   555   560

Pro Leu Ser Ala Val Asp Ala Glu Val Ser Arg His Leu Phe Glu Leu
            565   570   575

Cys Ile Cys Gln Ile Leu His Glu Lys Ile Thr Ile Leu Val Thr His
            580   585   590

Gln Leu Gln Tyr Leu Lys Ala Ala Ser Gln Ile Leu Ile Leu Lys Asp
    595   600   605

Gly Lys Met Val Gln Lys Gly Thr Tyr Thr Glu Phe Leu Lys Ser Gly
    610   615   620

Ile Asp Phe Gly Ser Leu Leu Lys Lys Asp Asn Glu Glu Ser Glu Gln

20

```
625                      630                      635                      640

Pro Pro Val Pro Gly Thr Pro Thr Leu Arg Asn Arg Thr Phe Ser Glu
            645                 650                 655

Ser Ser Val Trp Ser Gln Gln Ser Ser Arg Pro Ser Leu Lys Asp Gly
            660                 665                 670

Ala Leu Glu Ser Gln Asp Thr Glu Asn Val Pro Val Thr Leu Ser Glu
        675                 680                 685

Glu Asn Arg Ser Glu Gly Lys Val Gly Phe Gln Ala Tyr Lys Asn Tyr
    690                 695                 700

Phe Arg Ala Gly Ala His Trp Ile Val Phe Ile Phe Leu Ile Leu Leu
705             710                 715                 720

Asn Thr Ala Ala Gln Val Ala Tyr Val Leu Gln Asp Trp Trp Leu Ser
            725                 730                 735

Tyr Trp Ala Asn Lys Gln Ser Met Leu Asn Val Thr Val Asn Gly Gly
            740                 745                 750

Gly Asn Val Thr Glu Lys Leu Asp Leu Asn Trp Tyr Leu Gly Ile Tyr
        755                 760                 765

Ser Gly Leu Thr Val Ala Thr Val Leu Phe Gly Ile Ala Arg Ser Leu
    770                 775                 780

Leu Val Phe Tyr Val Leu Val Asn Ser Ser Gln Thr Leu His Asn Lys
785             790                 795                 800

Met Phe Glu Ser Ile Leu Lys Ala Pro Val Leu Phe Phe Asp Arg Asn
            805                 810                 815

Pro Ile Gly Arg Ile Leu Asn Arg Phe Ser Lys Asp Ile Gly His Leu
        820                 825                 830

Asp Asp Leu Leu Pro Leu Thr Phe Leu Asp Phe Ile Gln Thr Leu Leu
    835                 840                 845

Gln Val Val Gly Val Val Ser Val Ala Val Ala Val Ile Pro Trp Ile
    850                 855                 860

Ala Ile Pro Leu Val Pro Leu Gly Ile Ile Phe Ile Phe Leu Arg Arg
865             870                 875                 880

Tyr Phe Leu Glu Thr Ser Arg Asp Val Lys Arg Leu Glu Ser Thr Thr
            885                 890                 895

Arg Ser Pro Val Phe Ser His Leu Ser Ser Ser Leu Gln Gly Leu Trp
        900                 905                 910

Thr Ile Arg Ala Tyr Lys Ala Glu Glu Arg Cys Gln Glu Leu Phe Asp
        915                 920                 925

Ala His Gln Asp Leu His Ser Glu Ala Trp Phe Leu Phe Leu Thr Thr
    930                 935                 940
```

21

Ser Arg Trp Phe Ala Val Arg Leu Asp Ala Ile Cys Ala Met Phe Val
945   950   955   960

Ile Ile Val Ala Phe Gly Ser Leu Ile Leu Ala Lys Thr Leu Asp Ala
                965   970   975

Gly Gln Val Gly Leu Ala Leu Ser Tyr Ala Leu Thr Leu Met Gly Met
            980   985   990

Phe Gln Trp Cys Val Arg Gln Ser Ala Glu Val Glu Asn Met Met Ile
        995   1000   1005

Ser Val Glu Arg Val Ile Glu Tyr Thr Asp Leu Glu Lys Glu Ala
    1010   1015   1020

Pro Trp Glu Tyr Gln Lys Arg Pro Pro Pro Ala Trp Pro His Glu
    1025   1030   1035

Gly Val Ile Ile Phe Asp Asn Val Asn Phe Met Tyr Ser Pro Gly
    1040   1045   1050

Gly Pro Leu Val Leu Lys His Leu Thr Ala Leu Ile Lys Ser Gln
    1055   1060   1065

Glu Lys Val Gly Ile Val Gly Arg Thr Gly Ala Gly Lys Ser Ser
    1070   1075   1080

Leu Ile Ser Ala Leu Phe Arg Leu Ser Glu Pro Glu Gly Lys Ile
    1085   1090   1095

Trp Ile Asp Lys Ile Leu Thr Thr Glu Ile Gly Leu His Asp Leu
    1100   1105   1110

Arg Lys Lys Met Ser Ile Ile Pro Gln Glu Pro Val Leu Phe Thr
    1115   1120   1125

Gly Thr Met Arg Lys Asn Leu Asp Pro Phe Asn Glu His Thr Asp
    1130   1135   1140

Glu Glu Leu Trp Asn Ala Leu Gln Glu Val Gln Leu Lys Glu Thr
    1145   1150   1155

Ile Glu Asp Leu Pro Gly Lys Met Asp Thr Glu Leu Ala Glu Ser
    1160   1165   1170

Gly Ser Asn Phe Ser Val Gly Gln Arg Gln Leu Val Cys Leu Ala
    1175   1180   1185

Arg Ala Ile Leu Arg Lys Asn Gln Ile Leu Ile Ile Asp Glu Ala
    1190   1195   1200

Thr Ala Asn Val Asp Pro Arg Thr Asp Glu Leu Ile Gln Lys Lys
    1205   1210   1215

Ile Arg Glu Lys Phe Ala His Cys Thr Val Leu Thr Ile Ala His
    1220   1225   1230

Arg Leu Asn Thr Ile Ile Asp Ser Asp Lys Ile Met Val Leu Asp
    1235   1240   1245

22

Ser Gly Arg Leu Lys Glu Tyr Asp Glu Pro Tyr Val Leu Leu Gln
    1250   1255   1260

Asn Lys Glu Ser Leu Phe Tyr Lys Met Val Gln Gln Leu Gly Lys
    1265   1270   1275

Ala Glu Ala Ala Ala Leu Thr Glu Thr Ala Lys Gln Val Tyr Phe
    1280   1285   1290

Lys Arg Asn Tyr Pro His Ile Gly His Thr Asp His Met Val Thr
    1295   1300   1305

Asn Thr Ser Asn Gly Gln Pro Ser Thr Leu Thr Ile Phe Glu Thr
    1310   1315   1320

Ala Leu
    1325

<210>   22
<211>   332
<212>   PRT
<213>   Homo sapiens

<220>
<221>   misc_feature
<222>   (1)..(332)
<223>   X = any amino acid

<400>   22

Met Gly Met Phe Gln Trp Cys Val Arg Gln Ser Ala Glu Val Glu Asn
1   5   10   15

Met Met Ile Ser Val Glu Arg Val Ile Glu Tyr Thr Asp Leu Glu Lys
            20   25   30

Glu Ala Pro Trp Glu Tyr Gln Lys Arg Pro Pro Pro Ala Trp Pro His
        35   40   45

Glu Gly Val Ile Ile Phe Asp Asn Val Asn Phe Met Tyr Ser Pro Gly
    50   55   60

Gly Pro Leu Val Leu Lys His Leu Thr Ala Leu Ile Lys Ser Gln Glu
65   70   75   80

Lys Val Gly Ile Val Gly Arg Thr Gly Ala Gly Lys Ser Ser Leu Ile
            85   90   95

Ser Ala Leu Phe Arg Leu Ser Glu Pro Glu Gly Lys Ile Trp Ile Asp
        100   105   110

Lys Ile Leu Thr Thr Glu Ile Gly Leu His Asp Leu Arg Lys Lys Met
        115   120   125

Ser Ile Ile Pro Gln Glu Pro Val Leu Phe Thr Gly Thr Met Arg Lys
    130   135   140

```
Asn Leu Asp Pro Phe Asn Glu His Thr Asp Glu Glu Leu Trp Asn Ala
145             150             155             160

Leu Gln Glu Val Gln Leu Lys Glu Thr Ile Glu Asp Leu Pro Gly Lys
            165             170             175

Met Asp Thr Glu Leu Ala Glu Ser Gly Ser Asn Phe Ser Val Gly Gln
            180             185             190

Arg Gln Leu Val Cys Leu Ala Arg Ala Ile Leu Arg Lys Asn Gln Ile
        195             200             205

Leu Ile Ile Asp Glu Ala Thr Ala Asn Val Asp Pro Arg Thr Asp Glu
    210             215             220

Leu Ile Gln Lys Lys Ile Arg Glu Lys Phe Ala His Cys Thr Val Leu
225             230             235             240

Thr Ile Ala His Arg Leu Asn Thr Ile Ile Asp Ser Asp Lys Ile Met
            245             250             255

Val Leu Asp Ser Gly Arg Leu Lys Glu Tyr Asp Glu Pro Tyr Val Leu
            260             265             270

Leu Gln Asn Lys Glu Ser Leu Phe Phe Thr Arg Trp Val Ala Asn His
        275             280             285

Leu Gly Pro Gly Xaa Glu Ala Ala Cys Pro Ser Leu Gly Asn Xaa Pro
    290             295             300

Asn Arg Val Leu Leu Ser Gln Arg Asp Phe Xaa Pro Leu Phe Gly Phe
305             310             315             320

Thr Xaa Gly Pro Leu Gly Val Xaa Asn Pro Phe Pro
            325             330
```

```
<210>  23
<211>  954
<212>  PRT
<213>  Homo sapiens

<400>  23
Met Leu Pro Val Tyr Gln Glu Val Lys Pro Asn Pro Leu Gln Asp Ala
1           5               10              15

Asn Ile Cys Ser Arg Val Phe Phe Trp Trp Leu Asn Pro Leu Phe Lys
            20              25              30

Ile Gly His Lys Arg Arg Leu Glu Glu Asp Asp Met Tyr Ser Val Leu
        35              40              45

Pro Glu Asp Arg Ser Gln His Leu Gly Glu Glu Leu Gln Gly Phe Trp
    50              55              60

Asp Lys Glu Val Leu Arg Ala Glu Asn Asp Ala Gln Lys Pro Ser Leu
65              70              75              80

Thr Arg Ala Ile Ile Lys Cys Tyr Trp Lys Ser Tyr Leu Val Leu Gly
```

```
                    85                        90                        95

Ile Phe Thr Leu Ile Glu Glu Ser Ala Lys Val Ile Gln Pro Ile Phe
            100                   105                   110

Leu Gly Lys Ile Ile Asn Tyr Phe Glu Asn Tyr Asp Pro Met Asp Ser
            115               120                   125

Val Ala Leu Asn Thr Ala Tyr Ala Tyr Ala Thr Val Leu Thr Phe Cys
            130               135 ·                 140

Thr Leu Ile Leu Ala Ile Leu His His Leu Tyr Phe Tyr His Val Gln
145                   150               155                   160

Cys Ala Gly Met Arg Leu Arg Val Ala Met Cys His Met Ile Tyr Arg
                165               170               175

Lys Ala Leu Arg Leu Ser Asn Met Ala Met Gly Lys Thr Thr Thr Gly
            180               185               190

Gln Ile Val Asn Leu Leu Ser Asn Asp Val Asn Lys Phe Asp Gln Val
            195               200               205

Thr Val Phe Leu His Phe Leu Trp Ala Gly Pro Leu Gln Ala Ile Ala
    210               215               220

Val Thr Ala Leu Leu Trp Met Glu Ile Gly Ile Ser Cys Leu Ala Gly
225               230               235                   240

Met Ala Val Leu Ile Ile Leu Leu Pro Leu Gln Ser Cys Phe Gly Lys
            245               250               255

Leu Phe Ser Ser Leu Arg Ser Lys Thr Ala Thr Phe Thr Asp Ala Arg
            260               265               270

Ile Arg Thr Met Asn Glu Val Ile Thr Gly Ile Arg Ile Ile Lys Met
    275               280               285

Tyr Ala Trp Glu Lys Ser Phe Ser Asn Leu Ile Thr Asn Leu Arg Lys
    290               295               300

Lys Glu Ile Ser Lys Ile Leu Arg Ser Ser Cys Leu Arg Gly Met Asn
305               310               315                   320

Leu Ala Ser Phe Phe Ser Ala Ser Lys Ile Ile Val Phe Val Thr Phe
            325               330               335

Thr Thr Tyr Val Leu Leu Gly Ser Val Ile Thr Ala Ser Arg Val Phe
            340               345               350

Val Ala Val Thr Leu Tyr Gly Ala Val Arg Leu Thr Val Thr Leu Phe
            355               360               365

Phe Pro Ser Ala Ile Glu Arg Val Ser Glu Ala Ile Val Ser Ile Arg
    370               375               380

Arg Ile Gln Thr Phe Leu Leu Leu Asp Glu Ile Ser Gln Arg Asn Arg
385               390               395                   400
```

25

Gln Leu Pro Ser Asp Gly Lys Lys Met Val His Val Gln Asp Phe Thr
            405   410   415

Ala Phe Trp Asp Lys Ala Ser Glu Thr Pro Thr Leu Gln Gly Leu Ser
            420   425   430

Phe Thr Val Arg Pro Gly Glu Leu Leu Ala Val Val Gly Pro Val Gly
        435   440   445

Ala Gly Lys Ser Ser Leu Leu Ser Ala Val Leu Gly Glu Leu Ala Pro
    450   455   460

Ser His Gly Leu Val Ser Val His Gly Arg Ile Ala Tyr Val Ser Gln
465   470   475   480

Gln Pro Trp Val Phe Ser Gly Thr Leu Arg Ser Asn Ile Leu Phe Gly
            485   490   495

Lys Lys Tyr Glu Lys Glu Arg Tyr Glu Lys Val Ile Lys Ala Cys Ala
            500   505   510

Leu Lys Lys Asp Leu Gln Leu Leu Glu Asp Gly Asp Leu Thr Val Ile
    515   520   525

Gly Asp Arg Gly Thr Thr Leu Ser Gly Gly Gln Lys Ala Arg Val Asn
    530   535   540

Leu Ala Arg Ala Val Tyr Gln Asp Ala Asp Ile Tyr Leu Leu Asp Asp
545   550   555   560

Pro Leu Ser Ala Val Asp Ala Glu Val Ser Arg His Leu Phe Glu Leu
            565   570   575

Cys Ile Cys Gln Ile Leu His Glu Lys Ile Thr Ile Leu Val Thr His
            580   585   590

Gln Leu Gln Tyr Leu Lys Ala Ala Ser Gln Ile Leu Ile Leu Lys Asp
    595   600   605

Gly Lys Met Val Gln Lys Gly Thr Tyr Thr Glu Phe Leu Lys Ser Gly
    610   615   620

Ile Asp Phe Gly Ser Leu Leu Lys Lys Asp Asn Glu Glu Ser Glu Gln
625   630   635   640

Pro Pro Val Pro Gly Thr Pro Thr Leu Arg Asn Arg Thr Phe Ser Glu
            645   650   655

Ser Ser Val Trp Ser Gln Gln Ser Ser Arg Pro Ser Leu Lys Asp Gly
            660   665   670

Ala Leu Glu Ser Gln Asp Thr Glu Asn Val Pro Val Thr Leu Ser Glu
    675   680   685

Glu Asn Arg Ser Glu Gly Lys Val Gly Phe Gln Ala Tyr Lys Asn Tyr
    690   695   700

Phe Arg Ala Gly Ala His Trp Ile Val Phe Ile Phe Leu Ile Leu Leu
705   710   715   720

26

```
Asn Thr Ala Ala Gln Val Ala Tyr Val Leu Gln Asp Trp Trp Leu Ser
                 725             730                 735

Tyr Trp Ala Asn Lys Gln Ser Met Leu Asn Val Thr Val Asn Gly Gly
             740             745                 750

Gly Asn Val Thr Glu Lys Leu Asp Leu Asn Trp Tyr Leu Gly Ile Tyr
         755             760                 765

Ser Gly Leu Thr Val Ala Thr Val Leu Phe Gly Ile Ala Arg Ser Leu
     770             775             780

Leu Val Phe Tyr Val Leu Val Asn Ser Ser Gln Thr Leu His Asn Lys
785             790             795                 800

Met Phe Glu Ser Ile Leu Lys Ala Pro Val Leu Phe Phe Asp Arg Asn
             805             810                 815

Pro Ile Gly Arg Ile Leu Asn Arg Phe Ser Lys Asp Ile Gly His Leu
             820             825                 830

Asp Asp Leu Leu Pro Leu Thr Phe Leu Asp Phe Ile Gln Thr Leu Leu
         835             840             845

Gln Val Val Gly Val Val Ser Val Ala Val Ala Val Ile Pro Trp Ile
     850             855             860

Ala Ile Pro Leu Val Pro Leu Gly Ile Ile Phe Ile Phe Leu Arg Arg
865             870             875                 880

Tyr Phe Leu Glu Thr Ser Arg Asp Val Lys Arg Leu Glu Ser Thr Thr
             885             890                 895

Arg Ser Pro Val Phe Ser His Leu Ser Ser Ser Leu Gln Gly Leu Trp
             900             905                 910

Thr Ile Arg Ala Tyr Lys Ala Glu Glu Arg Cys Gln Glu Leu Phe Asp
         915             920             925

Ala His Gln Asp Leu His Ser Gly Leu Ser Ile Ser Gly Asn Gly Phe
     930             935             940

Lys Gly Gln Asp Leu Leu Leu Phe Leu Ala
945             950
```

```
<210> 24
<211> 301
<212> PRT
<213> Homo sapiens

<400> 24
Met Leu Pro Val Tyr Gln Glu Val Lys Pro Asn Pro Leu Gln Asp Ala
1               5                   10                  15

Asn Ile Cys Ser Arg Val Phe Phe Trp Trp Leu Asn Pro Leu Phe Lys
             20              25                  30
```

27

```
Ile Gly His Lys Arg Arg Leu Glu Glu Asp Asp Met Tyr Ser Val Leu
        35              40              45

Pro Glu Asp Arg Ser Gln His Leu Gly Glu Glu Leu Gln Gly Phe Trp
    50              55              60

Asp Lys Glu Val Leu Arg Ala Glu Asn Asp Ala Gln Lys Pro Ser Leu
65              70              75              80

Thr Arg Ala Ile Ile Lys Cys Tyr Trp Lys Ser Tyr Leu Val Leu Gly
            85              90              95

Ile Phe Thr Leu Ile Glu Glu Ser Ala Lys Val Ile Gln Pro Ile Phe
        100             105             110

Leu Gly Lys Ile Ile Asn Tyr Phe Glu Asn Tyr Asp Pro Met Asp Ser
        115             120             125

Val Ala Leu Asn Thr Ala Tyr Ala Tyr Ala Thr Val Leu Thr Phe Cys
    130             135             140

Thr Leu Ile Leu Ala Ile Leu His His Leu Tyr Phe Tyr His Val Gln
145             150             155             160

Cys Ala Gly Met Arg Leu Arg Val Ala Met Cys His Met Ile Tyr Arg
            165             170             175

Lys Ala Leu Arg Leu Ser Asn Met Ala Met Gly Lys Thr Thr Thr Gly
        180             185             190

Gln Ile Val Asn Leu Leu Ser Asn Asp Val Asn Lys Phe Asp Gln Val
        195             200             205

Thr Val Phe Leu His Phe Leu Trp Ala Gly Pro Leu Gln Ala Ile Ala
    210             215             220

Val Thr Ala Leu Leu Trp Met Glu Ile Gly Ile Ser Cys Leu Ala Gly
225             230             235             240

Met Ala Val Leu Ile Ile Leu Leu Pro Leu Gln Ser Cys Phe Gly Lys
            245             250             255

Leu Phe Ser Ser Leu Arg Ser Lys Thr Ala Thr Phe Thr Asp Ala Arg
        260             265             270

Ile Arg Thr Met Asn Glu Val Ile Thr Gly Ile Arg Ile Ile Lys Met
        275             280             285

Tyr Ala Cys Gly Lys Thr Gly Ser Ser Glu Ser Leu Ala
    290             295             300


<210>   25
<211>   336
<212>   PRT
<213>   Homo sapiens

<400>   25
Met Gly Met Phe Gln Trp Cys Val Arg Gln Ser Ala Glu Val Glu Asn
```

```
1                    5                          10                        15

Met Met Ile Ser Val Glu Arg Val Ile Glu Tyr Thr Asp Leu Glu Lys
            20                  25                  30

Glu Ala Pro Trp Glu Tyr Gln Lys Arg Pro Pro Pro Ala Trp Pro His
            35                  40                  45

Glu Gly Val Ile Ile Phe Asp Asn Val Asn Phe Met Tyr Ser Pro Gly
        50                  55                  60

Gly Pro Leu Val Leu Lys His Leu Thr Ala Leu Ile Lys Ser Gln Glu
65                  70                  75                  80

Lys Val Gly Ile Val Gly Arg Thr Gly Ala Gly Lys Ser Ser Leu Ile
                85                  90                  95

Ser Ala Leu Phe Arg Leu Ser Glu Pro Glu Gly Lys Ile Trp Ile Asp
            100                 105                 110

Lys Ile Leu Thr Thr Glu Ile Gly Leu His Asp Leu Arg Lys Lys Met
            115                 120                 125

Ser Ile Ile Pro Gln Glu Pro Val Leu Phe Thr Gly Thr Met Arg Lys
    130                 135                 140

Asn Leu Asp Pro Phe Asn Glu His Thr Asp Glu Glu Leu Trp Asn Ala
145                 150                 155                 160

Leu Gln Glu Val Gln Leu Lys Glu Thr Ile Glu Asp Leu Pro Gly Lys
                165                 170                 175

Met Asp Thr Glu Leu Ala Glu Ser Gly Ser Asn Phe Ser Val Gly Gln
            180                 185                 190

Arg Gln Leu Val Cys Leu Ala Arg Ala Ile Leu Arg Lys Asn Gln Ile
            195                 200                 205

Leu Ile Ile Asp Glu Ala Thr Ala Asn Val Asp Pro Arg Thr Asp Glu
    210                 215                 220

Leu Ile Gln Lys Lys Ile Arg Glu Lys Phe Ala His Cys Thr Val Leu
225                 230                 235                 240

Thr Ile Ala His Arg Leu Asn Thr Ile Ile Asp Ser Asp Lys Ile Met
                245                 250                 255

Val Leu Asp Ser Gly Arg Leu Lys Glu Tyr Asp Glu Pro Tyr Val Leu
            260                 265                 270

Leu Gln Asn Lys Glu Ser Leu Phe Tyr Lys Met Val Gln Gln Leu Gly
            275                 280                 285

Lys Ala Glu Ala Ala Ala Leu Thr Glu Thr Ala Lys Gln Val Tyr Phe
    290                 295                 300

Lys Arg Asn Tyr Pro His Ile Gly His Thr Asp His Met Val Thr Asn
305                 310                 315                 320
```

29

```
Thr Ser Asn Gly Gln Pro Ser Thr Leu Thr Ile Phe Glu Thr Ala Leu
            325                 330                 335


<210>  26
<211>  470
<212>  PRT
<213>  Homo sapiens

<400>  26
Met Lys Pro Pro Arg Arg Pro Ala Cys Ala Val Trp Trp Ile Arg Asp
1               5                   10                  15

His Gly Ala Trp Glu Ala Ala Gly Phe Ser Ala Pro Ser Leu Thr Glu
            20                  25                  30

Leu Phe Ser Leu Ser Val Ser Pro Ala Arg Ser Pro Val Phe Ser His
        35                  40                  45

Leu Ser Ser Ser Leu Gln Gly Leu Trp Thr Ile Arg Ala Tyr Lys Ala
    50                  55                  60

Glu Glu Arg Cys Gln Glu Leu Phe Asp Ala His Gln Asp Leu His Ser
65                  70                  75                  80

Glu Ala Trp Phe Leu Phe Leu Thr Thr Ser Arg Trp Phe Ala Val Arg
                85                  90                  95

Leu Asp Ala Ile Cys Ala Met Phe Val Ile Ile Val Ala Phe Gly Ser
            100                 105                 110

Leu Ile Leu Ala Lys Thr Leu Asp Ala Gly Gln Val Gly Leu Ala Leu
        115                 120                 125

Ser Tyr Ala Leu Thr Leu Met Gly Met Phe Gln Trp Cys Val Arg Gln
    130                 135                 140

Ser Ala Glu Val Glu Asn Met Met Ile Ser Val Glu Arg Val Ile Glu
145                 150                 155                 160

Tyr Thr Asp Leu Glu Lys Glu Ala Pro Trp Glu Tyr Gln Lys Arg Pro
                165                 170                 175

Pro Pro Ala Trp Pro His Glu Gly Val Ile Ile Phe Asp Asn Val Asn
            180                 185                 190

Phe Met Tyr Ser Pro Gly Gly Pro Leu Val Leu Lys His Leu Thr Ala
        195                 200                 205

Leu Ile Lys Ser Gln Glu Lys Val Gly Ile Val Gly Arg Thr Gly Ala
    210                 215                 220

Gly Lys Ser Ser Leu Ile Ser Ala Leu Phe Arg Leu Ser Glu Pro Glu
225                 230                 235                 240

Gly Lys Ile Trp Ile Asp Lys Ile Leu Thr Thr Glu Ile Gly Leu His
                245                 250                 255

Asp Leu Arg Lys Lys Met Ser Ile Ile Pro Gln Glu Pro Val Leu Phe
```

```
                    260                    265                    270

Thr Gly Thr Met Arg Lys Asn Leu Asp Pro Phe Asn Glu His Thr Asp
        275             280             285

Glu Glu Leu Trp Asn Ala Leu Gln Glu Val Gln Leu Lys Glu Thr Ile
    290             295             300

Glu Asp Leu Pro Gly Lys Met Asp Thr Glu Leu Ala Glu Ser Gly Ser
305             310             315             320

Asn Phe Ser Val Gly Gln Arg Gln Leu Val Cys Leu Ala Arg Ala Ile
            325             330             335

Leu Arg Lys Asn Gln Ile Leu Ile Ile Asp Glu Ala Thr Ala Asn Val
        340             345             350

Asp Pro Arg Thr Asp Glu Leu Ile Gln Lys Lys Ile Arg Glu Lys Phe
        355             360             365

Ala His Cys Thr Val Leu Thr Ile Ala His Arg Leu Asn Thr Ile Ile
    370             375             380

Asp Ser Asp Lys Ile Met Val Leu Asp Ser Gly Arg Leu Lys Glu Tyr
385             390             395             400

Asp Glu Pro Tyr Val Leu Leu Gln Asn Lys Glu Ser Leu Phe Tyr Lys
            405             410             415

Met Val Gln Gln Leu Gly Lys Ala Glu Ala Ala Ala Leu Thr Glu Thr
            420             425             430

Ala Lys Gln Val Tyr Phe Lys Arg Asn Tyr Pro His Ile Gly His Thr
        435             440             445

Asp His Met Val Thr Asn Thr Ser Asn Gly Gln Pro Ser Thr Leu Thr
    450             455             460

Ile Phe Glu Thr Ala Leu
465             470
```

CORRECTED VERSION

(19) World Intellectual Property
Organization
International Bureau

(43) International Publication Date
18 December 2003 (18.12.2003)

PCT

(10) International Publication Number
WO 2003/104404 A2

(51) International Patent Classification⁷: C12N

(21) International Application Number:
PCT/US2003/017772

(22) International Filing Date: 5 June 2003 (05.06.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/386,651 6 June 2002 (06.06.2002) US

(71) Applicant (for all designated States except US): AVALON PHARMACEUTICALS, INC. [US/US]; 20358 Seneca Meadows Parkway, Germantown, MD 20876 (US).

(72) Inventor; and
(75) Inventor/Applicant (for US only): EBNER, Reinhard [US/US]; 9906 Shelburne Terrace #316, Gaithersburg, MD 20878 (US).

(74) Agents: GRANT, Alan, J. et al.; Bain, Gilfillan, Cecchi, Stewart & Olstein, 6 Becker Farm Road, Roseland, NJ 07068 (US).

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— without international search report and to be republished upon receipt of that report

(48) Date of publication of this corrected version:
21 May 2004

(15) Information about Correction:
see PCT Gazette No. 21/2004 of 21 May 2004, Section II

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: CANCER-LINKED GENE AS TARGET FOR CHEMOTHERAPY

(57) Abstract: Cancer-linked gene sequences, and derived amino acid sequences, are disclosed along with processes for assaying potential antitumor agents based on their modulation of the expression of these cancer-linked genes. Also disclosed are antibodies that react with the disclosed polypeptides and methods of using the antibodies to treat cancerous conditions, such as by using the antibody to target cancerous cells in vivo for purposes of delivering therapeutic agents thereto. Also described are methods of diagnosing using the gene sequences as well as the polypeptide sequences, especially for diagnosing prostate cancer.